

POISSON AND COMPOUND POISSON APPROXIMATIONS IN A NONCONVENTIONAL SETUP

YURI KIFER AND ARIEL RAPAPORT

INSTITUTE OF MATHEMATICS
HEBREW UNIVERSITY
JERUSALEM, ISRAEL

ABSTRACT. It was shown in [17] that for any subshift of finite type considered with a Gibbs invariant measure the numbers of multiple recurrences to shrinking cylindrical neighborhoods of almost all points are asymptotically Poisson distributed. Here we not only extend this result to all ψ -mixing shifts with countable alphabet but actually show that for all points the distributions of these numbers are asymptotically close either to Poisson or to compound Poisson distributions. It turns out that for all nonperiodic points a limiting distribution is always Poisson while at the same time for periodic points there may be no limiting distribution at all unless the shift invariant measure is Bernoulli in which case the limiting distribution always exists. As a corollary we obtain also that the first occurrence time of the multiple recurrence event is asymptotically exponentially distributed. Most of the results are new also for widely studied single recurrences (see, for instance, [14], [15], [3] and [5]), as well.

1. INTRODUCTION

Ergodic theorems for nonconventional sums of the form

$$(1.1) \quad S_N = S_N(\omega) = \sum_{k=1}^N f_1(T^{q_1(k)}\omega) f_2(T^{q_2(k)}\omega) \cdots f_\ell(T^{q_\ell(k)}\omega),$$

were initiated in [10] and employed there in the proof of Szemerédi's theorem on arithmetic progressions while the name "nonconventional" comes from [11]. Here T is an ergodic measure preserving transformation, f_i 's are bounded measurable functions and $q_i(k) = ik$, $i = 1, \dots, \ell$. Since then results concerning such ergodic theorems under various conditions evolved into a substantial body of literature. More recently under appropriate mixing conditions a strong law of large numbers and a functional central limit theorem were obtained even for more general sums in [16] and [18], respectively.

If f_i equals for each i an indicator \mathbb{I}_A of the same measurable set A then the corresponding sum $S_N = S_N^A(\omega)$ counts the multiple recurrence events $T^{q_i(k)}\omega \in$

Date: December 11, 2012.

2000 Mathematics Subject Classification. Primary: 60F05 Secondary: 37D35, 60J05.

Key words and phrases. Poisson and compound Poisson approximations, nonconventional sums, ψ -mixing, subshift of finite type.

A , $i = 1, \dots, \ell$ which occur for $k \leq N$. It was shown in [17] that if we count such multiple arrivals to appropriately shrinking sets A_n then the sums

$$(1.2) \quad S_N^{A_n}(\omega) = \sum_{k=1}^N \mathbb{I}_{A_n}(T^{q_1(k)}\omega) \mathbb{I}_{A_n}(T^{q_2(k)}\omega) \cdots \mathbb{I}_{A_n}(T^{q_\ell(k)}\omega)$$

usually will have asymptotically Poisson distribution for suitably chosen sequences $N = N_n \rightarrow \infty$ as $n \rightarrow \infty$. For $\ell = 1$ and $q_1(k) = k$ this type of results was obtained in a series of papers under various conditions (see, for instance, [14], [15], [3] and [5]).

To explain the results of this paper more precisely let us specify first our setup which consists of a stationary ψ -mixing discrete time process $\xi(k)$, $k = 0, 1, \dots$ evolving on a finite or countable state space \mathcal{A} and of nonnegative increasing functions q_i , $i = 1, \dots, \ell$ taking on integer values on integers and such that $q_1(k) < q_2(k) < \dots < q_\ell(k)$ when $k \geq 1$. For each sequence $a = (a_0, a_1, \dots) \in \mathcal{A}^\mathbb{N}$ of elements from \mathcal{A} and any $n \in \mathbb{N}$ denote by $a^{(n)}$ the string a_0, a_1, \dots, a_{n-1} which determines also an n -cylinder set A_n^a in $\mathcal{A}^\mathbb{N}$ consisting of sequences whose initial n -string coincides with a_0, a_1, \dots, a_{n-1} . For appropriately chosen sequences $N = N_n$ we are interested in the number of those $l \leq N$ such that the process $\xi(k) = \xi(k, \omega)$, $k \geq 0$ repeats the string $a^{(n)}$ starting at times $q_1(l), q_2(l), \dots, q_\ell(l)$. Employing the left shift transformation T on the sequence space $\mathcal{A}^\mathbb{N}$ we can represent the number in question via $S_N^A(\omega)$ given by (1.2) with $A = A_n^a$ and $N = N_n$ considering S_N^A as a random variable on the probability space corresponding to the process ξ .

Viewing such $S_N^{A_n^a}$ as random variables on $\mathcal{A}^\mathbb{N}$ considered with a Gibbs T -invariant measure \mathbb{P} it was shown in [17] that $S_N^{A_n^a}$ for almost all $a \in \mathcal{A}^\mathbb{N}$ has asymptotically a Poisson distribution with a parameter t provided

$$(1.3) \quad N = N_n = N_n^a \sim t(\mathbb{P}(A_n^a))^{-\ell} \quad \text{as } n \rightarrow \infty.$$

Observe that such asymptotic in n results make sense only for sequences $a \in \mathcal{A}^\mathbb{N}$ such that $\mathbb{P}(A_n^a) > 0$ for all $n \geq 1$ but this is good enough since the set of such a 's has probability one. We will show in this paper for linear q_i 's that under (1.3) for large n the distribution of $S_N^{A_n^a}$ with N as in (1.3) is close either to a Poisson distribution with a parameter t or to a compound Poisson distribution. The latter or the former holds true depending on whether or not the string $a^{(n)}$ looks periodic with relatively short with respect to n period. In the "conventional" case $\ell = 1$ Poisson approximation estimates were obtained in [4] and [5] while compound Poisson approximation estimates for periodic sequences were derived in [15] but even in this case the complete dichotomy as described above seems to be new.

Relying on our approximation estimates it is possible to see (assuming (1.3)) for which sequences a the distribution of $S_N^{A_n^a}$ approaches for large n the Poisson distribution and for which compound Poisson. Moreover, we will show that under (1.3) for all nonperiodic sequences a the sum $S_N^{A_n^a}$ converges in distribution as $n \rightarrow \infty$ to a Poisson random variable with the parameter t . On the other hand, for periodic sequences a the limiting distribution may not exist at all and we provide a corresponding example. We observe that an attempt to construct such an example was made in Section 3.4 of [15] but it follows both from our estimates and, actually, already from [5] that for the example in [15] the limiting distribution

exists and it is Poisson. We prove also that if \mathbb{P} is a product stationary measure on $\mathcal{A}^{\mathbb{N}}$ (Bernoulli measure), i.e. if coordinate projections are independent identically distributed (i.i.d.) random variables, then the limiting distributions, either Poissonian or compound Poissonian, exist for all sequences a . The above results describe completely the limiting behavior as $n \rightarrow \infty$ in this setup and they seem to be new even for the extensively studied $\ell = 1$ case. Furthermore, either of the above convergence results ruins also the example in [15] since the latter counts arrivals to cylinders constructed by a nonperiodic sequence and it is built on a shift space with a product probability measure. Nevertheless, as our nonconvergence example shows a slight perturbation of independency may already create sequences a where convergence fails.

Let $a \in \mathcal{A}^{\mathbb{N}}$ and $\tau_{A_n^a}(\omega)$ be the first time k when the string $a^{(n)}$ starts at each of the places $q_1(k), q_2(k), \dots, q_\ell(k)$ of a sequence $\omega \in \mathcal{A}^{\mathbb{N}}$. When $\ell = 1$ and $q_1(k) = k$ such hitting times were studied in a number of papers (see, for instance, [2], [3] and references there) where it was shown that under appropriate normalization they have as $n \rightarrow \infty$ exponential limiting distribution. As a corollary of our Poisson and compound Poisson approximations we will extend here this type of results to the nonconventional $\ell > 1$ situation.

Our results are applicable to larger classes of dynamical systems and not only to shifts. Indeed, any expansive endomorphism S of a compact M , in particular, any smooth expanding endomorphism of a compact manifolds, has a symbolic representation as a one sided shift by taking a finite partition $\alpha_1, \dots, \alpha_k$ of M into sets of small diameter and assigning to each point $x \in M$ the sequence j_0, j_1, \dots such that $x \in \cap_{i=0}^{\infty} S^{-i} \alpha_{j_i}$ while noticing that the last intersection is a singelton in view of expansivity. Smooth expanding endomorphisms of compact manifolds have many exponentially fast ψ -mixing invariant measures which are Gibbs measures constructed by Hölder continuous functions which ensures fast decay of our approximation estimates in this case. Our results remain valid also for many invertible dynamical systems having symbolic representations as two sided shifts, notably, for Axiom A (in particular, Anosov) diffeomorphisms. Within number theoretic applications the results can be formulated in terms of occurrence of prescribed strings of digits in base- m or continued fraction expansions.

The structure of this paper is the following. In the next section we describe precisely our setup and conditions, give necessary definitions, formulate our main results and discuss more their connections and relevance. In Section 3 we state and prove some auxiliary lemmas. In Sections 4 and 5 our Poisson and compound Poisson approximations results will be proved. In Section 6 we prove existence of limiting distributions in the i.i.d. case while in Section 7 we exhibit our nonconvergence example.

2. PRELIMINARIES AND MAIN RESULTS

We start with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that Ω is a space of sequences $\mathcal{A}^{\mathbb{N}}$ with entries from a finite or countable set (alphabet) \mathcal{A} which is not a singelton, the σ -algebra \mathcal{F} is generated by cylinder sets and a probability \mathbb{P} invariant with respect to the left shift T acting by $(T\omega)_i = \omega_{i+1}$ for $\omega = (\omega_0, \omega_1, \dots) \in \Omega = \mathcal{A}^{\mathbb{N}}$. For each $I \subset \mathbb{N}$ denote by \mathcal{F}_I the sub σ -algebra of \mathcal{F} generated by the cylinder sets $[a_i, i \in I] = \{\omega = (\omega_0, \omega_1, \dots) \in \Omega : \omega_i = a_i \forall i \in I\}$. Without loss of generality we assume that the probability of each 1-cylinder set is positive, i.e. $\mathbb{P}([a]) > 0$ for

every $a \in \mathcal{A}$, and since \mathcal{A} is not a singelton we have also $\sup_{a \in \mathcal{A}} \mathbb{P}([a]) < 1$. Our results will be based on the ψ -mixing (dependence) coefficient defined for any two σ -algebras $\mathcal{G}, \mathcal{H} \subset \mathcal{F}$ by (see [9]),

$$(2.1) \quad \begin{aligned} \psi(\mathcal{G}, \mathcal{H}) &= \sup_{B \in \mathcal{G}, C \in \mathcal{H}} \left\{ \left| \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(B)\mathbb{P}(C)} - 1 \right|, \mathbb{P}(B)\mathbb{P}(C) \neq 0 \right\} \\ &= \sup \{ \|\mathbb{E}(g|\mathcal{G}) - Eg\|_\infty : g \text{ is } \mathcal{H} - \text{measurable and } \mathbb{E}|g| \leq 1 \} \end{aligned}$$

where $\|\cdot\|_p$ is the $L^p(\Omega, \mathbb{P})$ -norm. Next, we set

$$(2.2) \quad \psi_m = \sup_{n \in \mathbb{N}} \psi(\mathcal{F}_{0,n}, \mathcal{F}_{n+m+1,\infty})$$

where $\mathcal{F}_{k,n} = \mathcal{F}_{\{i: k \leq i \leq n\}}$ when $0 \leq k \leq n \leq \infty$. It follows from (2.1) and (2.2) that ψ_m is non increasing in m and the measure \mathbb{P} is called ψ -mixing if $\psi_0 < \infty$ and $\psi_m \rightarrow 0$ as $m \rightarrow \infty$.

Another important ingredient of our setup is a collection of ℓ positive increasing integer valued functions q_1, \dots, q_ℓ defined on positive integers \mathbb{N}_+ and such that

$$(2.3) \quad \begin{aligned} q_1(n) &< q_2(n) < \dots < q_\ell(n), \forall n \in \mathbb{N}_+ \text{ and} \\ \lim_{n \rightarrow \infty} (q_{i+1}(n) - q_i(n)) &= \infty, \forall i = 1, \dots, \ell - 1. \end{aligned}$$

Our estimates will involve the following quantities related to these functions

$$(2.4) \quad g(n) = \inf_{k \geq n} \min_{1 \leq i \leq \ell-1} (q_{i+1}(k) - q_i(k)) \text{ and } \gamma(n) = \min\{k \geq 0 : g(k) \geq 2n\}$$

where $n \in \mathbb{N}_+$. In view of the second assumption in (2.3) the function $\gamma(n)$ is well defined for all n .

Denote by \mathcal{C}_n the set of all n -cylinders $A = [a_0, , a_1, \dots, a_{n-1}]$. Here and in what follows we denote as above by $[Q]$ a cylinder if Q is a finite string of elements from \mathcal{A} but, as usual, when Q is a number then $[Q]$ denotes the integral part of Q . Similarly to [5] we introduce also for each $A = [a_0, a_1, \dots, a_{n-1}] \in \mathcal{C}_n$ the quantity

$$(2.5) \quad \pi(A) = \min\{k \in \{1, 2, \dots, n\} : A \cap T^{-k}A \neq \emptyset\}$$

setting also $A(\pi) = [a_0, a_1, \dots, a_{\pi(A)-1}] \in \mathcal{C}_{\pi(A)}$. Next, for each $A \in \mathcal{C}_n$ we write

$$(2.6) \quad S_N^A = \sum_{k=1}^N X_k^A \text{ where } X_k^A = \prod_{i=1}^{\ell} \mathbb{I}_A \circ T^{q_i(k)}$$

and often, when A will be fixed and clear, we will drop the index A in X_k^A and S_N^A . In order to shorten our formulas we will use throughout this paper the notation $\wp(x) = xe^x$. The following result exhibits our Poisson approximation estimates (for $\ell = 1$ a similar estimate was obtained in [5]).

2.1. Theorem. *Let $A \in \mathcal{C}_n$ with $\mathbb{P}(A) > 0$, $t > 0$, $N = [t(\mathbb{P}(A))^{-\ell}]$ and assume that $\psi_n < (3/2)^{1/(\ell+1)} - 1$. Then*

$$(2.7) \quad \begin{aligned} \sup_{L \subset \mathbb{N}} |\mathbb{P}\{S_N^A \in L\} - P_t(L)| &\leq 16\mathbb{P}(A)(\ell^2 nt + \gamma(n)(1 + t^{-1})) \\ &\quad + 6\mathbb{P}(A(\pi))tn\ell^2(1 + \psi_0) + 2\wp(2^\ell t\psi_n + \gamma(n)\mathbb{P}(A)) \end{aligned}$$

where $P_t(L)$ is the probability assigned to L by the Poisson distribution with the parameter t . Moreover, assume that \mathbb{P} is ψ -mixing (i.e. $\psi_n \rightarrow 0$ as $n \rightarrow \infty$) then (2.7) holds true with

$$(2.8) \quad \mathbb{P}(A) \leq e^{-\Gamma n} \text{ and } \mathbb{P}(A(\pi)) \leq e^{-\Gamma \pi(A)}$$

for some $\Gamma > 0$ independent of A and n .

Next, for each $\omega = (\omega_0, \omega_1, \dots) \in \Omega$ and $n \geq 1$ set $A_n^\omega = [\omega_0, \dots, \omega_{n-1}]$. Denote by $\Omega_{\mathbb{P}}$ the set of $\omega \in \Omega$ such that $\mathbb{P}(A_n^\omega) > 0$ for all $n \geq 1$. Clearly, $T\Omega_{\mathbb{P}} \subset \Omega_{\mathbb{P}}$ and $\mathbb{P}(\Omega_{\mathbb{P}}) = 1$ since $\Omega_{\mathbb{P}}$ is the complement in Ω of the union of all cylinder sets $A \in \mathcal{C}_n$, $n \geq 1$ such that $\mathbb{P}(A) = 0$ and the number of such cylinders is countable. Theorem 2.1 yields the following asymptotic result.

2.2. Corollary. *Assume that the conditions of Theorem 2.1 together with the ψ -mixing assumption hold true. Then for \mathbb{P} -almost all $\omega \in \Omega$ there exists $M(\omega) < \infty$ such that for any $n \geq M(\omega)$,*

$$(2.9) \quad \sup_{L \subset \mathbb{N}} |\mathbb{P}\{S_N^{A_n^\omega} \in L\} - P_t(L)| \leq 16e^{-\Gamma n} (\ell^2 nt + \gamma(n)(1 + t^{-1}) + t\ell^2 n^4(1 + \psi_0)) + 2\wp(2^\ell t\psi_n + \gamma(n)e^{-\Gamma n})$$

provided $N = N_n^\omega = [t(\mathbb{P}(A_n^\omega))^{-\ell}]$.

Theorem 2.1 yields that if both $\pi(A_n^\omega)$ and $g(n)$ grow fast enough in n then under the ψ -mixing condition the distribution of S_N^A approaches the Poisson one as $n \rightarrow \infty$. The fast growth of $\pi(A_n^\omega)$ in n means that ω is "very" nonperiodic and Corollary 2.2 sais that almost all ω 's fall into this category. The above results can be compared with Theorem 2.3 from [17] where \mathbb{P} was supposed to be a Gibbs invariant measure corresponding to a Hölder continuous function concentrated on a subshift of finite type space (see [8]). A ψ -mixing coefficient of such a measure decays in n exponentially fast (see [8]), and so (2.8) holds true then. Under conditions of [17] the function $g(n)$ grows faster than logarithmically, and so $\gamma(n)$ grows slower than exponentially which yields a fast decay of the errors in the Poisson approximations above which improves the result from [17] where no estimates of the convergence rate were obtained. Note also that any Gibbs measure \mathbb{P} gives a positive weight to each cylinder in the corresponding subshift of finite type space, and so in this case the latter space coincides with $\Omega_{\mathbb{P}}$. We observe that in the "conventional" $\ell = 1$ case similar to Theorem 2.1 error estimates in the corresponding Poisson approximations were obtained in [4] and [5].

Next, we describe our compound Poisson approximations where we assume that

$$(2.10) \quad q_i(k) = d_i k \text{ for some } d_i \in \mathbb{N}, i = 1, \dots, \ell \text{ with } 1 \leq d_1 < d_2 < \dots < d_\ell.$$

For each $R = [a_0, a_1, \dots, a_{r-1}] \in \mathcal{C}_r$ and an integer $n \geq r$ set

$$R^{n/r} = (\cap_{k=0}^{[n/r]-1} T^{-kr}(R)) \cap T^{-[n/r]}([a_0, \dots, a_{n-r[n/r]-1}])$$

where we define $[a_0, \dots, a_{n-r[n/r]-1}] = \Omega$ if r divides n . Observe that if $A = [a_0, a_1, \dots, a_{n-1}]$, $A(\pi) = R$ and $\pi(A) = r$ then $R^{n/r} = A$.

2.3. Theorem. *Let $A = [a_0, a_1, \dots, a_{n-1}] \in \mathcal{C}_n$ with $\mathbb{P}(A) > 0$, $t > 0$, $N = [t(\mathbb{P}(A))^{-\ell}]$ and assume that $\psi_n < (3/2)^{1/(\ell+1)} - 1$. Set $r = \pi(A)$, $R = [a_0, \dots, a_{r-1}]$,*

$$(2.11) \quad \kappa = lcm\left\{\frac{r}{gcd\{r, d_i\}} : 1 \leq i \leq \ell\right\} \text{ and } \rho = \rho_A = \prod_{i=1}^{\ell} \mathbb{P}\{R^{(n+d_i\kappa)/r} | A\}$$

where lcm and gcd denote the least common multiple and the greatest common divisor, respectively. Assume that \mathbb{P} is ψ -mixing then

$$(2.12) \quad \sup_{A \in \mathcal{C}_n, n \geq 1} \rho_A < 1$$

and if $\omega \in \Omega$ is not a periodic sequence then

$$(2.13) \quad \lim_{n \rightarrow \infty} \rho_{A_n^\omega} = 0$$

where A_n^ω is the same as in Corollary 2.2. Furthermore, let W be a Poisson random variable with the parameter $t(1 - \rho)$. Then for any $n > r(d_\ell + 6)$ there exists a sequence of i.i.d. random variables η_1, η_2, \dots independent of W such that $\mathbb{P}\{\eta_1 \in \{1, \dots, [\frac{n}{r}]\}\} = 1$ and the compound Poisson random variable $Z = \sum_{k=1}^W \eta_k$ satisfies

$$(2.14) \quad \sup_{L \subset \mathbb{N}} |\mathbb{P}\{S_N^A \in L\} - \mathbb{P}\{Z \in L\}| \leq 2^{2\ell+7}(1 + \psi_0)^{2\ell}(d_\ell \ell^2 n^4 e^{-\Gamma n/2} + \psi_n(1 - e^{-\Gamma})^{-1}) + 2\wp(10(1 + \psi_0)^{2\ell} d_\ell n^2(t + 1)e^{-\Gamma n/2} + 2^\ell t \psi_n).$$

We will see in Lemma 3.7 that for any nonperiodic sequence $\omega \in \Omega$ both (2.13) and $r_n^\omega = \pi(A_n^\omega) \rightarrow \infty$ as $n \rightarrow \infty$ hold true which together with appropriate estimates for the distribution of η_i 's constructed in Theorem 2.3 will yield the following result.

2.4. Corollary. *Under conditions and notations of Theorem 2.3 for any nonperiodic sequence $\omega \in \Omega_{\mathbb{P}}$ the sum $S_N^{A_n^\omega}$ converges in distribution to a Poisson random variable with the parameter t .*

An important role in the proof of estimates in both Theorems 2.1 and 2.3 will play the estimates from [6]. We observe that both Theorem 2.3 and Corollary 2.4 seem to be new even for the "conventional" case $\ell = 1$. Some compound Poisson approximation estimates were obtained in [15] but they deal there only with cylinders $A = A_n^a$ constructed for periodic sequences a and only with geometrically distributed random variables η_i appearing in the compound Poisson random variable Z . Only results about convergence for almost all sequences ω were known before while Corollary 2.4 asserts convergence for all but periodic points whose number is countable.

2.5. Remark. In fact, all above results can be formulated assuming that $N \sim t(\mathbb{P}(A))^{-\ell}$ instead of $N = [t(\mathbb{P}(A))^{-\ell}]$ where $A \in \mathcal{C}_n$. Indeed, in view of Lemma 3.2 of the next section for all sufficiently large (in comparison to n) k ,

$$\mathbb{E}X_k^A \leq C(P(A))^\ell$$

for some $C > 0$, and so

$$\sum_{\min(N, t(\mathbb{P}(A))^{-\ell}) \leq k \leq \max(N, t(\mathbb{P}(A))^{-\ell})} X_k^A \rightarrow 0 \text{ in probability as } n \rightarrow \infty.$$

Next, for any $A \in \mathcal{C}_n$ define

$$(2.15) \quad \tau_A(\omega) = \min\{k \geq 1 : \prod_{i=1}^{\ell} \mathbb{I}_A \circ T^{q_i(k)}(\omega) = 1\}$$

which is the first time k for which the multiple recurrence event $\{T^{q_i(k)}\omega \in A, i = 1, \dots, \ell\}$ occurs. Then combining Theorems 2.1 and 2.3 (assuming zero occurrence of multi recurrence events in question) we derive

2.6. Corollary. *Suppose that \mathbb{P} is ψ -mixing and $q_i(k) = d_i k$, $i = 1, \dots, \ell$. Then for any $A \in \mathcal{C}_n$ with $\mathbb{P}(A) > 0$ and $t \geq 0$,*

$$(2.16) \quad |\mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > t\} - e^{-(1-\rho_A)t}| \leq 2^{2\ell+8}(1 + \psi_0)^{2\ell}(t + 1)(\psi_n(1 - e^{-\Gamma})^{-1} + d_\ell \ell^2 n^4(1 + t^{-1})e^{-\Gamma n/(d_\ell + 6)}) + 2\wp(2^\ell t \psi_n + 10e^{-\Gamma n/2}(1 + \psi_0)^{2\ell} d_\ell n^2(t + 1)),$$

provided $\psi_n < (3/2)^{1/(\ell+1)} - 1$.

We observe that when $\ell = 1$ the random variable τ_A can be treated as a stopping time and most direct methods which deal with τ_A in this case are based on this fact (see, for instance, [2], [3] and references there). In our nonconventional situation when $\ell > 1$ the random variable τ_A depends on the future and it is difficult to deal with it directly. By this reason we obtain Corollary 2.6 as an immediate consequence of Poisson and compound Poisson approximations applied to the case when no multiple recurrence event occurs until time $t(\mathbb{P}(A))^{-\ell}$. Observe that if we replace t in (2.16) by $u = t(1 - \rho_A)$ then in view of (2.12) Corollary 2.6 will take the form of Theorem 1 from [2] which dealt though only with the "conventional" $\ell = 1$ case. Recall, also that arrivals to more general than cylindrical sets may lead to other limiting distribution (see [20]) so in order to remain in the class of Poisson and compound Poisson distributions we have to restrict the investigation to arrivals to cylinder sets.

In Section 5 we will derive from Corollary 2.6 together with the Shannon-McMillan-Breiman theorem the following result.

2.7. Corollary. *Suppose that (2.10) holds true and the ψ -mixing coefficient of \mathbb{P} satisfies*

$$(2.17) \quad \sum_{n=1}^{\infty} \psi_n \ln n < \infty.$$

Then for \mathbb{P} -almost all $\omega \in \Omega$,

$$(2.18) \quad \lim_{n \rightarrow \infty} \frac{1}{n} (\ln \tau_{A_n^\omega} + \ell \ln \mathbb{P}(A_n^\omega)) = 0 \quad \mathbb{P} - a.s.$$

If, in addition,

$$(2.19) \quad - \sum_{a \in \mathcal{A}} \mathbb{P}([a]) \ln \mathbb{P}([a]) < \infty$$

then for \mathbb{P} -almost all $\omega \in \Omega$,

$$(2.20) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \ln \tau_{A_n^\omega} = h_{\mathbb{P}}(T) \quad \mathbb{P} - a.s.$$

where $h_{\mathbb{P}}(T)$ is the Kolmogorov-Sinai entropy of the shift T with respect to its invariant measure \mathbb{P} .

In the "conventional" $\ell = 1$ case relying on estimates of Theorem 1 from [2] it is possible to obtain (2.20) under weaker than (2.17) conditions.

Under additional independency conditions we obtain a more specific than in Theorem 2.3 compound Poisson approximation of S_n^A constructed via independent geometrically distributed random variables.

2.8. Theorem. *Suppose that the conditions of Theorem 2.3 are satisfied and assume in addition that \mathbb{P} is the product stationary probability on $\mathcal{A}^{\mathbb{N}}$, i.e. that coordinate projections from Ω to \mathcal{A} are i.i.d. random variables. Let W be the same as in Theorem 2.3 and ζ_1, ζ_2, \dots be i.i.d. random variables independent of W and having the geometric distribution with the parameter ρ defined in Theorem 2.3, i.e. $\mathbb{P}\{\zeta_1 = k\} = (1 - \rho)\rho^{k-1}$, $k = 1, 2, \dots$. Let $Y = \sum_{k=1}^W \zeta_k$ then*

$$(2.21) \quad \begin{aligned} & \sup_{L \subset \mathbb{N}} |\mathbb{P}\{S_N^A \in L\} - \mathbb{P}\{Y \in L\}| \\ & \leq 2^{2\ell+8} (t+1) \ell^2 d_\ell n^4 e^{-\Gamma n/2} + 2\wp(12d_\ell n^2(t+1)e^{-\Gamma n/2}) + \frac{t^{[n/r]+1}}{([n/r]+1)!}. \end{aligned}$$

where A , N , and r are the same as in Theorem 2.3.

We observe that when \mathbb{P} is a product stationary measure on $\Omega = \mathcal{A}^{\mathbb{N}}$ then $\Omega_{\mathbb{P}} = \Omega$ since then in view of our assumption that $\mathbb{P}([a]) > 0$ for all $a \in \mathcal{A}$ any cylinder set has positive probability. The following result, which seems to be new even in the "conventional" $\ell = 1$ case, sais that under the independency conditions of Theorem 2.8 the limiting distribution, either Poisson or compound Poisson, always exists.

2.9. Theorem. *Suppose that the conditions of Theorem 2.8 are satisfied. For each $\omega \in \Omega$ and $n \geq 1$ set $U_n^\omega = S_{N_n^\omega}^{A_n^\omega}$ where A_n^ω is the same as in Corollary 2.2, $N_n^\omega = [t(\mathbb{P}(A_n^\omega))^{-\ell}]$ and $t > 0$. Let $r_n^\omega = \pi(A_n^\omega)$, $R_n^\omega = [\omega_0, \omega_1, \dots, \omega_{r_n^\omega-1}]$ and $\rho_n^\omega = \mathbb{P}(R_n^\omega)$. Then the limits*

$$(2.22) \quad \lim_{n \rightarrow \infty} r_n^\omega = r^\omega \text{ and } \lim_{n \rightarrow \infty} \rho_n^\omega = \rho^\omega$$

exist for any $\omega \in \Omega$.

(i) If ω is not a periodic point then U_n^ω converges in distribution as $n \rightarrow \infty$ to a Poisson random variable with the parameter t .

(ii) If ω is a periodic point (sequence) of a minimal period d then $r^\omega = d$ and $\rho^\omega = (\mathbb{P}(A_d^\omega))^{k_0}$ where

$$k_0 = \frac{\kappa^\omega}{r^\omega} \sum_{i=1}^{\ell} d_i$$

with κ^ω defined by (2.11) for r^ω in place of r there. Furthermore, let W be a Poisson random variable with the parameter $t(1 - \rho^\omega)$ and ζ_1, ζ_2, \dots be a sequence of i.i.d. random variables independent of W having the geometric distribution with the parameter ρ^ω , i.e. $\mathbb{P}\{\zeta_1 = k\} = (1 - \rho^\omega)(\rho^\omega)^{k-1}$, $k = 1, 2, \dots$. Then U_n^ω converges in distribution to $Y^\omega = \sum_{k=1}^W \zeta_k$ as $n \rightarrow \infty$.

A number theory (combinatorial) application of Theorem 2.9 can be described in the following way. Fix a point $a \in [0, 1)$ having a base $m \geq 2$ expansion $a = \sum_{i=0}^{\infty} a_i m^{-(i+1)}$. For each point $\omega \in [0, 1)$ with a base m expansion $\omega = \sum_{i=0}^{\infty} \omega_i m^{-(i+1)}$ count the number $N_n(\omega)$ of those $i \leq tm^n$ for which the initial n -string a_0, a_1, \dots, a_{n-1} of a is repeated starting from all places $d_1 i, d_2 i, \dots, d_\ell i$ in the expansion of ω . Then the random variable $N_n(\omega)$ on the probability space $([0, 1], \text{Leb})$ converges in distribution either to a Poisson (with the parameter t) or to a compound Poisson random variable depending on whether the expansion of a is nonperiodic or periodic. In particular, for all irrational and for some rational a the limiting distribution will be Poissonian.

Since we allow also countable alphabet shift spaces our results are applicable to continued fraction expansions, as well. There the Gauss map $Gx = \frac{1}{x} \pmod{1}$, $x \in (0, 1]$, $G0 = 0$ preserves the Gauss measure $\mu_G(\Gamma) = \frac{1}{\ln 2} \int_{\Gamma} \frac{dx}{1+x}$ which is exponentially fast ψ -mixing (see, for instance, [13]) and G (restricted to points with infinite continued fraction expansions) is conjugate to the left shift on a sequence space with a countable alphabet. Thus Theorems 2.1 and 2.3 are applicable here, as well. On the other hand, Theorem 2.9 does not work for the Gauss measure which does not produce i.i.d. digits in continued fraction expansions (see, for instance, [19]), and so in order to apply Theorem 2.9 here we have to take other probability measures conjugated to product measures on the shift space which produce i.i.d. continued fraction digits. Moreover, these results remain valid for a larger class of transformations generated by, so called, f -expansions (see [13]).

It follows from both Corollary 2.4 and Theorem 2.9 that the nonconvergence example from Section 3.4 in [15] is not correct and, in fact, the limiting distribution exists there and it is Poisson since the cylinder sets constructed there were based on a nonperiodic point and the probability measure on the sequence space was a product measure, whence both Corollary 2.4 and Theorem 2.9 above are applicable. Still, we construct in Section 7 an example of a point $\omega \in \Omega$ and of a ψ -mixing shift invariant measure such that $S_{N_n^\omega}^{A_n^\omega}$ considered with $\ell = 1$ and $d_1 = 1$ does not have a limiting distribution as $n \rightarrow \infty$. Moreover, the ψ -mixing coefficient ψ_l in our example equals zero for any $l \geq 1$ while $\psi_0 < \infty$, i.e. the situation there is as close as possible to independence where Theorem 2.9 asserts convergence for all points $\omega \in \Omega$.

2.10. Remark. Observe that our convergence theorems above are based on approximation estimates of Theorems 2.1 and 2.3 which yield convergence of distributions with respect to the total variation distance but for integer valued random variables we are dealing with this is equivalent to convergence in distribution.

2.11. Remark. All above results were stated for one-sided shifts but essentially without changes they remain true for two-sided shifts, as well. We can also restrict the discussion to a subshift of finite type space considering invariant probability measures supported there. This enables us to apply the results to Axiom A diffeomorphisms (see [8]) considered with their Gibbs invariant measures relying on their symbolic representations via Markov partitions.

3. AUXILIARY LEMMAS

We start with the following result.

3.1. Lemma. *Suppose that \mathbb{P} is ψ -mixing then there exists a constant $\Gamma > 0$ such that for any $A \in \mathcal{C}_n$,*

$$(3.1) \quad \mathbb{P}(A) \leq e^{-\Gamma n}.$$

Proof. The proof is contained in Lemma 1 from [1] and in Lemma 1 from [12]. In both places the authors assume summability of ψ_n but, in fact, both proofs use only convergence of ψ_n to zero as $n \rightarrow \infty$ which is ψ -mixing. In both papers the result is formulated with a constant in front of the exponent in (3.1) but it is clear from the proofs there that this constant can be taken equal 1. Furthermore, the authors there work on finite alphabet shift spaces but the proof of this result remains valid without any changes for countable alphabets, as well. \square

We will need also the following result which is, essentially, well known but for readers' convenience we give its simple proof here.

3.2. Lemma. *Let Q_1, \dots, Q_l , $l \geq 1$ be subsets of nonnegative integers such that for $i = 2, \dots, l - 1$,*

$$(3.2) \quad \max Q_{i-1} + k \leq \min Q_i \leq \max Q_i \leq \min Q_{i+1} - k$$

for some integer $k > 0$. Then for any $U_i \in \mathcal{F}_{Q_i}$, $i = 1, \dots, l$,

$$(3.3) \quad |\mathbb{P}(\cap_{i=1}^l U_i) - \prod_{i=1}^l \mathbb{P}(U_i)| \leq ((1 + \psi_k)^l - 1) \prod_{i=1}^l \mathbb{P}(U_i).$$

Proof. By the definition (2.1) and (2.2) of the coefficient ψ ,

$$(3.4) \quad |\mathbb{P}(\cap_{i=1}^{j+1} U_i) - \mathbb{P}(\cap_{i=1}^j U_i)\mathbb{P}(U_{j+1})| \leq \psi_k \mathbb{P}(\cap_{i=1}^j U_i)\mathbb{P}(U_{j+1})$$

for $j = 1, \dots, l-1$. applying (3.4) successively $l-1$ times we obtain

$$(3.5) \quad |\mathbb{P}(\cap_{i=1}^l U_i) - \prod_{i=1}^l \mathbb{P}(U_i)| \leq \psi_k \sum_{j=1}^{l-1} (\mathbb{P}(\cap_{i=1}^j U_i) \prod_{i=j+1}^l \mathbb{P}(U_i)).$$

Furthermore, applying (3.4) successively $j-1$ times we see that

$$(3.6) \quad \mathbb{P}(\cap_{i=1}^j U_i) \leq (1 + \psi_k) \mathbb{P}(\cap_{i=1}^{j-1} U_i) \mathbb{P}(U_j) \leq (1 + \psi_k)^{j-1} \prod_{i=1}^j \mathbb{P}(U_i).$$

This together with (3.5) yields (3.3). \square

3.3. Lemma. let Q and \tilde{Q} be two subsets of nonnegative integers such that at least one of these sets is finite and

$$d = \min_{i \in Q, j \in \tilde{Q}} |i - j| > 0.$$

Let k bound the number of components of Q and \tilde{Q} which are separated by some elements of the other set. Assume that $\psi_d < 2^{1/k} - 1$. Then

$$(3.7) \quad \psi(\mathcal{F}_Q, \mathcal{F}_{\tilde{Q}}) \leq 2^{2k+2} \psi_d (2 - (1 + \psi_d)^k)^{-2}.$$

Proof. Suppose, for instance, that Q is a finite set. Then Q and \tilde{Q} can be represented as disjoint unions

$$Q = \cup_{i=1}^k Q_{2i-1} \text{ and } \tilde{Q} = \cup_{i=1}^k Q_{2i}$$

such that Q_1 and Q_{2k} may be empty sets while all Q_j , $j = 2, \dots, 2k-1$ are nonempty and

$$(3.8) \quad \max Q_{j-1} + d \leq \min Q_j \leq \max Q_j \leq \min Q_{j+1} - d$$

for $j = 2, \dots, 2k-1$ where if $j = 2$ and $Q_1 = \emptyset$ or if $j = 2k-1$ and $Q_{2k} = \emptyset$ then we disregard the first or the last inequality in (3.8), respectively.

Next, let $U = \cap_{i=1}^k U_{2i-1}$ and $V = \cap_{i=1}^k U_{2i}$ where $U_{2i-1} \in \mathcal{F}_{Q_{2i-1}}$ and $U_{2i} \in \mathcal{F}_{Q_{2i}}$, $i = 1, \dots, k$. Then by Lemma 3.2,

$$(3.9) \quad |\mathbb{P}(U \cap V) - \prod_{j=1}^{2k} \mathbb{P}(U_j)| \leq ((1 + \psi_d)^{2k} - 1) \prod_{j=1}^{2k} \mathbb{P}(U_j),$$

$$(3.10) \quad |\mathbb{P}(U) - \prod_{i=1}^k \mathbb{P}(U_{2i-1})| \leq ((1 + \psi_d)^k - 1) \prod_{i=1}^k \mathbb{P}(U_{2i-1})$$

and

$$(3.11) \quad |\mathbb{P}(V) - \prod_{i=1}^k \mathbb{P}(U_{2i})| \leq ((1 + \psi_d)^k - 1) \prod_{i=1}^k \mathbb{P}(U_{2i}).$$

Combining (3.9)–(3.11) we obtain that

$$(3.12) \quad |\mathbb{P}(U \cap V) - \mathbb{P}(U)\mathbb{P}(V)| \leq ((1 + \psi_d)^{2k} - 1) \prod_{i=1}^k \mathbb{P}(U_{2i-1}) (\mathbb{P}(V) + 2 \prod_{i=1}^k \mathbb{P}(U_{2i})).$$

Next, by Lemma 3.2,

$$(3.13) \quad \mathbb{P}(U) \geq (2 - (1 + \psi_d)^k) \prod_{i=1}^k \mathbb{P}(U_{2i-1}) \text{ and } \mathbb{P}(V) \geq (2 - (1 + \psi_d)^k) \prod_{i=1}^k \mathbb{P}(U_{2i}),$$

which together with (3.12) yields that

$$(3.14) \quad \begin{aligned} & |\mathbb{P}(U \cap V) - \mathbb{P}(U)\mathbb{P}(V)| \\ & \leq ((1 + \psi_d)^{2k} - 1)(4 - (1 + \psi_d)^k)(2 - (1 + \psi_d)^k)^{-2}\mathbb{P}(U)\mathbb{P}(V). \end{aligned}$$

The inequality (3.14) remains true when U and V are disjoint unions of intersections of sets as above and it will still hold under monotone limits of sets $U \in \mathcal{F}_Q$ and $V \in \mathcal{F}_{\tilde{Q}}$. Hence, it holds true for all $U \in \mathcal{F}_Q$ and $V \in \mathcal{F}_{\tilde{Q}}$, and so (3.7) follows (see [7]). \square

We will need also the following estimate of the total variation distance between two Poisson distributions.

3.4. Lemma. *For any $\lambda, \gamma > 0$,*

$$(3.15) \quad \sum_{l=0}^{\infty} |P_{\lambda}(l) - P_{\gamma}(l)| \leq 2e^{|\lambda - \gamma|}|\lambda - \gamma| = 2\wp(|\lambda - \gamma|).$$

Proof. Assume, for instance, that $\lambda \geq \gamma$. Then

$$\begin{aligned} \sum_{n=0}^{\infty} |e^{-\lambda} \frac{\lambda^n}{n!} - e^{-\gamma} \frac{\gamma^n}{n!}| & \leq \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} (|1 - e^{\lambda - \gamma}| \\ & + e^{\lambda - \gamma} |1 - (\frac{\gamma}{\lambda})^n|) \leq |1 - e^{\lambda - \gamma}| + e^{\lambda - \gamma} |\lambda - \gamma| \leq 2(\lambda - \gamma)e^{\lambda - \gamma} \end{aligned}$$

and (3.15) follows. \square

The following two lemmas will be used in the proof of Theorem 2.3.

3.5. Lemma. *Let $H = [a_0, \dots, a_{h-1}] \in \mathcal{C}_h$, $h \geq 1$ and either $\pi(H) = h$ or h is not divisible by $\pi(H)$. Then $\pi(H^{n/h}) = h$ for each $n \geq 2h$.*

Proof. For each $B \in \mathcal{C}_m$ set

$$\mathcal{O}(B) = \{k \leq m, k \geq 1 : B \cap T^{-k}B \neq \emptyset\},$$

so that $\pi(B) = \min(\mathcal{O}(B))$. It follows from Theorem 6 and Remark 7 from [22] that we can also write

$$(3.16) \quad \begin{aligned} \mathcal{O}(B) &= \{\pi(B), 2\pi(B), \dots, [\frac{m}{\pi(B)}]\pi(B)\} \\ &\cup \{k \in \{m - \pi(B) + 1, \dots, m\} : B \cap T^{-k}B \neq \emptyset\}. \end{aligned}$$

Let $n \geq 2h$ and set $\pi(H^{n/h}) = k$. Clearly, $h \in \mathcal{O}(H^{n/h})$ and so $h \geq k$. We suppose that $h > k$ and arrive at a contradiction. By the definition of π it follows that $H^{n/h} \cap T^{-k}(H^{n/h}) \neq \emptyset$. Therefore, $[a_k, \dots, a_{h-1}] = [a_0, \dots, a_{h-1-k}]$ and $[a_0, \dots, a_{k-1}] = [a_{h-k}, \dots, a_{h-1}]$, and so $k, h - k \in \mathcal{O}(H)$. Thus, $k \geq \pi(H)$ and if $\pi(H) = h$ we would have $k \geq h$ which contradicts our assumption. Hence, it remains to consider the case when h is not divisible by $\pi(H)$. Since $k, h - k \in \mathcal{O}(H)$ then $\pi(H) \leq k, h - k$, and so $k, h - k \leq h - \pi(H)$. This together with (3.16) yields that there exist integers a and b such that $1 \leq a, b \leq [\frac{h}{\pi(H)}]$, $k = a \cdot \pi(H)$ and $h - k = b \cdot \pi(H)$. Therefore $h = (h - k) + k = (a + b)\pi(H)$ which contradicts our assumption that h is not divisible by $\pi(H)$. Hence, $h = k$ and the proof is complete. \square

The following result will be used in the proof of Theorem 2.3.

3.6. Lemma. *Let q_i , $i = 1, \dots, \ell$ be as in (2.10) and $n \geq r(d_\ell + 1)$. For any positive integers m and n satisfying $m > 2d_\ell n$ set*

$$\mathcal{N} = \mathcal{N}_{m,n} = \{l = 1, 2, \dots, n : \{X_m = 1, X_{m+l} = 1\} \neq \emptyset\}.$$

Then $\min \mathcal{N} = \kappa$ where κ is defined by (2.11) with n being the length of a cylinder A there and $r = \pi(A)$.

Proof. It follows from (3.16) that if $1 \leq l \leq \frac{n-r}{d_\ell}$ then $l \in \mathcal{N}$ if and only if r divides $d_i l$ for each $i = 1, \dots, \ell$. By the definition r divides κd_i and by the assumption of the lemma $\kappa \leq r \leq \frac{n-r}{d_\ell}$. Hence, $\kappa \in \mathcal{N}$, and so $\kappa \geq y = \min \mathcal{N}$. Now, let $l \in \mathcal{N}$ satisfies $l \leq \frac{n-r}{d_\ell}$. Then r divides $d_i l$ for each $1 \leq i \leq \ell$, and so $\frac{r}{\gcd\{r, d_i\}}$ divides l . Thus, κ divides l , and so $\kappa \leq l$. It follows that $\kappa \leq y$, completing the proof of the lemma. \square

In Sections 5 and 6 we will need the following results. As in Corollary 2.2 for each $\omega = (\omega_0, \omega_1, \dots) \in \Omega$ set $A_n^\omega = [\omega_0, \dots, \omega_{n-1}]$, $r_n^\omega = \pi(A_n^\omega)$ and $R_n^\omega = [\omega_0, \dots, \omega_{r_n^\omega-1}]$. Next, define κ_n^ω and ρ_n^ω by (2.11) with $r = r_n^\omega$ and $A = A_n^\omega$.

3.7. Lemma. *Given $\omega \in \Omega$ the limit $r^\omega = \lim_{n \rightarrow \infty} r_n^\omega$ exist. Furthermore, if ω is a periodic point with period $d \in \mathbb{N}^+$ (i.e. the whole path $\{T^k \omega : k \geq 0\}$ of ω consists of d points) then $r^\omega = d$, otherwise $r^\omega = \infty$. Assume that \mathbb{P} is ψ -mixing. If $\omega \in \Omega_{\mathbb{P}}$ is not a periodic point then*

$$(3.17) \quad \lim_{n \rightarrow \infty} \rho_n^\omega = 0.$$

Furthermore, if independency conditions of Theorem 2.9 are satisfied then the limit $\rho^\omega = \lim_{n \rightarrow \infty} \rho_n^\omega$ always exists. Moreover, in this case if ω is a periodic point with period $d \in \mathbb{N}^+$ then $\rho^\omega = (\mathbb{P}([\omega_0, \dots, \omega_{d-1}]))^{k_0}$ with k_0 defined in Theorem 2.9(ii), otherwise $\rho^\omega = 0$.

Proof. Assume that ω is periodic with period $d \in \mathbb{N}^+$ and set $D = [\omega_0, \dots, \omega_{d-1}]$. By the definition of a period of a point it follows that $\pi(D) = d$ or $\pi(D)$ does not divide d . This together with Lemma 3.5 yields that $\pi(D^{n/d}) = d$ for all $n \geq 2d$. Moreover, $T^d(\omega) = \omega$ which implies that $D^{n/d} = A_n^\omega$ for each $n \geq d$. From this it follows that

$$\lim_{n \rightarrow \infty} r_n^\omega = \lim_{n \rightarrow \infty} \pi(A_n^\omega) = \lim_{n \rightarrow \infty} \pi(D^{n/d}) = d$$

Now assume that ω is not periodic. Given $n \geq 1$, from the definition of \mathcal{O} in Lemma 3.5 it follows that $r_{n+1}^\omega \in \mathcal{O}(A_n^\omega) \cup \{n+1\}$, which implies that $r_n^\omega \leq r_{n+1}^\omega$. This holds true for all $n \geq 1$, and so $r^\omega = \lim_{n \rightarrow \infty} r_n^\omega$ exists and it is in the set $\mathbb{N}^+ \cup \{\infty\}$.

Assume by contradiction that $r^\omega < \infty$, then there exists an integer $M \geq 1$ such that $r_n^\omega = r^\omega$ for all $n \geq M$. From this it follows that $A_n^\omega = [\omega_0, \dots, \omega_{r^\omega-1}]^{n/r^\omega}$ for all such n which implies that $\omega = [\omega_0, \dots, \omega_{r^\omega-1}]^\infty$, and so ω is a periodic point which is a contradiction to our assumption. From this it follows that $r^\omega = \infty$ and the assertion concerning r^ω is proved.

Next, let $\omega \in \Omega_{\mathbb{P}}$ be not a periodic point. Let $r = r_n^\omega$, $\kappa = \kappa_n^\omega$, $A = A_n^\omega$ and $R = R_n^\omega$. Since r divides $d_i \kappa$ for each $i = 1, \dots, \ell$ we see by (2.1), (2.2) and (3.1) that

$$(3.18) \quad \mathbb{P}(R^{(n+d_i \kappa)/r}) \leq (1 + \psi_0) \mathbb{P}(A) \mathbb{P}(T^{n_r} R^{(n_r+r)/r}) \leq (1 + \psi_0) e^{-\Gamma r_n^\omega} \mathbb{P}(A)$$

where $n_r = n \pmod{r} = n - [n/r]r$. Now (3.17) follows since by the above $r_n^\omega \rightarrow \infty$ as $n \rightarrow \infty$ when ω is not periodic. Under the conditions of Theorem 2.9 the remaining assertion concerning ρ^ω follows from the properties of r_n^ω derived above together with the independence assumption. \square

The assertion (2.12) of Theorem 2.3 we obtain as a separate lemma.

3.8. Lemma. *Assume that \mathbb{P} is ψ -mixing then (2.12) holds true.*

Proof. First, clearly $\rho_A \leq 1$ for any $A \in \mathcal{C}_n$ and each $n \geq 1$. Next, let $A_n = [a_0^{(n)}, \dots, a_{n-1}^{(n)}] \in \mathcal{C}_n$ and write

$$(3.19) \quad \mathbb{P}(R_n^{(n+d_1\kappa_n)/r_n} | A_n) = \frac{\mathbb{P}(R_n^{(n+d_1\kappa_n)/r_n})}{\mathbb{P}(A_n)} = 1 - \delta_n, \quad \delta_n \geq 0$$

where $R_n = [a_0^{(n)}, \dots, a_{r_n-1}^{(n)}]$, $r_n = \pi(A_n)$ and κ_n is given by (2.11) with $r = r_n$. Then we show by induction that for any $k \in \mathbb{N}$,

$$(3.20) \quad \mathbb{P}(R_n^{(n+kd_1\kappa_n)/r_n}) \geq (1 - 2^{k-1}\delta_n)\mathbb{P}(A_n).$$

Indeed, (3.20) is satisfied for $k = 1$ in view of (3.19). Suppose that (3.20) holds true for $k = m$ and prove it for $k = m + 1$. Set $n_r = n - r_n[n/r_n]$ and observe that for any $l \geq 1$,

$$(3.21) \quad \begin{aligned} & (R_n^{(n+(l-1)d_1\kappa_n)/r_n} \cap T^{-(n+(l-1)d_1\kappa_n)}(\Omega \setminus T^{n_r} R_n^{(n_r+d_1\kappa_n)/r_n})) \\ & \cup R_n^{(n+ld_1\kappa_n)/r_n} = R_n^{(n+(l-1)d_1\kappa_n)/r_n} \subset A_n \end{aligned}$$

and the union above is disjoint. The induction hypothesis together with (3.21) considered with $l = m$ yields

$$(3.22) \quad \begin{aligned} & \mathbb{P}(R_n^{(n+md_1\kappa_n)/r_n} \cap T^{-(n+md_1\kappa_n)}(\Omega \setminus T^{n_r} R_n^{(n_r+d_1\kappa_n)/r_n})) \\ & \leq \mathbb{P}(T^{-d_1\kappa_n}(R_n^{(n+(m-1)d_1\kappa_n)/r_n} \cap T^{-(n+(m-1)d_1\kappa_n)}(\Omega \setminus T^{n_r} R_n^{(n_r+d_1\kappa_n)/r_n}))) \\ & = \mathbb{P}(R_n^{(n+(m-1)d_1\kappa_n)/r_n} \cap T^{-(n+(m-1)d_1\kappa_n)}(\Omega \setminus T^{n_r} R_n^{(n_r+d_1\kappa_n)/r_n})) \\ & \leq \delta_n 2^{m-1} \mathbb{P}(A_n). \end{aligned}$$

Employing (3.21) with $l = m + 1$ we obtain from (3.22) and (3.20) for $k = m$ that

$$\mathbb{P}(R_n^{(n+(m+1)d_1\kappa_n)/r_n}) \geq \mathbb{P}(R_n^{(n+md_1\kappa_n)/r_n}) - \delta_n 2^{m-1} \mathbb{P}(A_n) \geq (1 - 2^m \delta_n) \mathbb{P}(A_n),$$

and so (3.20) holds true with $k = m + 1$ completing the induction.

Now observe that by (2.1), (2.2) and (3.1),

$$(3.23) \quad \mathbb{P}(R_n^{(n+kd_1\kappa_n)/r_n}) \leq (1 + \psi_0)e^{-\Gamma kr_n} \mathbb{P}(A_n).$$

Since always $r_n \geq 1$ we can choose k so large that $(1 + \psi_0)e^{-\Gamma kr_n} < \frac{1}{2}$ for all n making the right hand side of (3.23) less than $\frac{1}{2}\mathbb{P}(A_n)$ for all n . Now suppose by contradiction that there exists a subsequence $l_m \rightarrow \infty$ as $m \rightarrow \infty$ such that $\delta_{l_m} \rightarrow 0$ as $m \rightarrow \infty$. Then, we can choose $n = l_m$ in (3.20) with m so large that the right hand side of (3.20) will be bigger than $\frac{1}{2}\mathbb{P}(A_n)$ which leads to a contradiction proving the lemma. \square

4. POISSON APPROXIMATION

4.1. Proof of Theorem 2.1. If $t^{-1}\gamma(n)(\mathbb{P}(A))^\ell \geq \frac{1}{4}$ then the theorem clearly holds true, and so we can assume that $\gamma(n) < \frac{t(\mathbb{P}(A))^{-\ell}}{4}$. Set $U = S_N^A$. If $t(\mathbb{P}(A))^{-\ell} < 4$ then for each $L \subset \mathbb{N}$,

$$(4.1) \quad \begin{aligned} |\mathbb{P}\{U \in L\} - P_t(L)| &\leq \mathbb{P}\{U \neq 0\} + |\mathbb{P}\{U = 0\} - P_t\{0\}| + P_t(\mathbb{N}^+) \\ &\leq 2\mathbb{P}\{U \neq 0\} + 2P_t(\mathbb{N}^+) \leq 2N\mathbb{P}(A) + 2(1 - e^{-t}) \leq 8\mathbb{P}(A) + 2t \leq 16\mathbb{P}(A). \end{aligned}$$

Again the theorem holds true, so we can assume that $t(\mathbb{P}(A))^{-\ell} \geq 4$ and then $\gamma(n) < \frac{t(\mathbb{P}(A))^{-\ell}}{4} \leq N$.

Set $W = \sum_{\alpha=\gamma(n)}^N X_\alpha$, where $X_\alpha = X_\alpha^A$ was defined in (2.6), and $\lambda = EW$. For any $L \subset \mathbb{N}$,

$$(4.2) \quad \begin{aligned} |\mathbb{P}\{U \in L\} - P_t(L)| &\leq |\mathbb{P}\{U \in L\} - \mathbb{P}\{W \in L\}| \\ &\quad + |\mathbb{P}\{W \in L\} - P_\lambda(L)| + |P_\lambda(L) - P_t(L)| = \delta_1 + \delta_2 + \delta_3 \end{aligned}$$

where δ_1, δ_2 and δ_3 denote the first, the second and the third terms in the right hand side of (4.2), respectively. We estimate δ_1 by

$$(4.3) \quad \delta_1 \leq 2\mathbb{P}\{U - W > 0\} \leq 2 \sum_{\alpha=1}^{\gamma(n)} \mathbb{P}\{X_\alpha = 1\} \leq 2\gamma(n)\mathbb{P}(A).$$

In order to estimate δ_2 we use Theorem 1 from [6]. Note that from the assumption $\psi_n \leq (3/2)^{1/(\ell+1)} - 1$ and from Lemma 3.2 whenever $\gamma(n) \leq \alpha \leq N$ it follows that

$$\mathbb{P}\{X_\alpha = 1\} \geq (2 - (1 + \psi_n)^\ell)(\mathbb{P}(A))^\ell > \frac{1}{2}(\mathbb{P}(A))^\ell > 0.$$

Hence, the conditions of Theorem 1 from [6] are satisfied with the collection $\{X_{\gamma(n)}, \dots, X_N\}$. For each $\alpha \in \mathbb{N}_+$ satisfying $\gamma(n) \leq \alpha \leq N$ set

$$B_\alpha = \{\beta \geq \gamma(n), \beta \leq N : \text{such that } |q_i(\alpha) - q_j(\beta)| < 2n \text{ for some } i, j = 1, 2, \dots, \ell\}.$$

Then by Theorem 1 from [6],

$$(4.4) \quad \delta_2 \leq b_1 + b_2 + b_3$$

where

$$b_1 = \sum_{\alpha=\gamma(n)}^N \left(\sum_{\beta \in B_\alpha} \mathbb{P}\{X_\alpha = 1\} \mathbb{P}\{X_\beta = 1\} \right),$$

$$b_2 = \sum_{\alpha=\gamma(n)}^N \left(\sum_{\alpha \neq \beta \in B_\alpha} \mathbb{P}\{X_\alpha = 1, X_\beta = 1\} \right) \text{ and } b_3 = \sum_{\alpha=\gamma(n)}^N E|E(X_\alpha - EX_\alpha | \mathcal{B}_\alpha)|$$

with $\mathcal{B}_\alpha = \sigma\{X_\beta : \beta \notin B_\alpha\}$.

From Lemma 3.2 and the fact that $|B_\alpha| \leq 4\ell^2 n$ for each α it follows that

$$(4.5) \quad b_1 \leq N4\ell^2 n ((1 + \psi_n)\mathbb{P}(A))^{2\ell} \leq 8\ell^2 nt(\mathbb{P}(A))^\ell.$$

Next, we estimate b_2 . Let $\gamma(n) \leq \alpha \leq N$ and $\alpha \neq \beta \in B_\alpha$. Assume without loss of generality that $\alpha < \beta$, so $q_1(\alpha) < q_1(\beta)$. If $q_1(\beta) - q_1(\alpha) < \pi(A)$ then by the

definition of $\pi(A)$ it follows that $\mathbb{P}\{X_\alpha = 1, X_\beta = 1\} = 0$, so we can assume that $q_1(\beta) - q_1(\alpha) \geq \pi(A)$. Now by (2.1), (2.2) and Lemmas 3.1 and 3.2,

$$\begin{aligned} \mathbb{P}\{X_\alpha = 1, X_\beta = 1\} &\leq \mathbb{P}(T^{-q_1(\alpha)}([a_0, \dots, a_{\pi(A)-1}]) \cap \{X_\beta = 1\}) \\ &\leq (1 + \psi_0)\mathbb{P}(A(\pi))\mathbb{P}\{X_\beta = 1\} \leq (1 + \psi_0)\mathbb{P}(A(\pi))(1 + \psi_n)^\ell(\mathbb{P}(A))^\ell. \end{aligned}$$

Since by our assumption $(1 + \psi_n)^\ell \leq 3/2$ we obtain that

$$(4.6) \quad b_2 \leq 6(1 + \psi_0)N\ell^2n\mathbb{P}(A(\pi))(\mathbb{P}(A))^\ell \leq 6(1 + \psi_0)\ell^2tn\mathbb{P}(A(\pi)).$$

In order to estimate b_3 we use Lemma 3.3. Fix an integer α such that $\gamma(n) \leq \alpha \leq N$ and set

$$\begin{aligned} Q &= Q_\alpha = \{q_i(\alpha) + m : i = 1, \dots, \ell; m = 0, 1, \dots, n-1\} \text{ and} \\ \tilde{Q} &= \tilde{Q}_\alpha = \{q_j(\beta) + m : j = 1, \dots, \ell; \beta \notin B_\alpha, m = 0, 1, \dots, n-1\}. \end{aligned}$$

Then the conditions of Lemma 3.3 are satisfied with $d = n$ and such Q and \tilde{Q} . Taking into account that $B_\alpha \subset \mathcal{F}_{\tilde{Q}}$ we derive easily from Lemmas 3.2 and 3.3 that for $p = EX_\alpha$,

$$\begin{aligned} E|E(X_\alpha - p | \mathcal{B}_\alpha)| &= E|E(E(X_\alpha - p | \mathcal{F}_{\tilde{Q}}) | \mathcal{B}_\alpha)| \\ &\leq EE(|E(X_\alpha - p | \mathcal{F}_{\tilde{Q}})| | \mathcal{B}_\alpha) \\ &= E|E(X_\alpha - p | \mathcal{F}_{\tilde{Q}})| \leq 2^{2\ell+4}\psi_n\mathbb{P}\{X_\alpha = 1\} \leq 2^{2\ell+5}\psi_n(\mathbb{P}(A))^\ell \end{aligned}$$

Hence,

$$(4.7) \quad b_3 \leq N2^{2\ell+5}\psi_n(\mathbb{P}(A))^\ell \leq 2^{2\ell+5}t\psi_n.$$

In order to estimate δ_3 we use Lemma 3.4 which yields

$$\delta_3 \leq \sum_{l=0}^{\infty} |P_\lambda\{l\} - P_t\{l\}| \leq 2e^{|\lambda-t|}|\lambda-t| = 2\wp(|\lambda-t|).$$

We also have by Lemma 3.2 that

$$\begin{aligned} |\lambda - t| &\leq |E(\sum_{\alpha=\gamma(n)}^N X_\alpha) - N(\mathbb{P}(A))^\ell| + (\mathbb{P}(A))^\ell \\ &\leq \sum_{\alpha=\gamma(n)}^N |\mathbb{P}\{X_\alpha = 1\} - (\mathbb{P}(A))^\ell| + \gamma(n)(\mathbb{P}(A))^\ell \\ &\leq N\psi_n 2^\ell(\mathbb{P}(A))^\ell + \gamma(n)(\mathbb{P}(A))^\ell \leq 2^\ell t\psi_n + \gamma(n)(\mathbb{P}(A))^\ell. \end{aligned}$$

It follows that

$$(4.8) \quad \delta_3 \leq 2\wp(2^\ell t\psi_n + \gamma(n)(\mathbb{P}(A))^\ell).$$

Now (2.7) follows from (4.1)–(4.8) while (2.8) follows from Lemma 3.1, completing the proof of the theorem. \square

4.2. Proof of Corollary 2.2. Set $c = 3\Gamma^{-1}$ and fix $M \in \mathbb{N}$ such that $M > c \ln M$ and $\psi_n \leq (3/2)^{1/(\ell+1)} - 1$ for all $n \geq M$. Denote by Ω^* the set of all $\omega \in \Omega$ for which there exist an $M(\omega) \geq M$ such that for each $n \geq M(\omega)$,

$$\pi(A_n^\omega) > n - c \ln n \text{ and } \mathbb{P}(A_n^\omega) > 0.$$

Set $U_n^\omega = S_N^{A_n^\omega}$. Assuming (2.8) it follows from Theorem 2.1 that for each $\omega \in \Omega^*$ and $n \geq M(\omega)$,

$$\begin{aligned} & \sup_{L \subset \mathbb{N}} |\mathbb{P}\{U_n^\omega \in L\} - P_t(L)| \\ & \leq 16e^{-\Gamma n} (\ell^2 nt + \gamma(n)(1+t^{-1}) + t\ell^2 n^4(1+\psi_0)) + 2\varphi(2^\ell t\psi_n + \gamma(n)e^{-\Gamma n}) \end{aligned}$$

which gives (2.11) and it remains to show that Ω^* has the full measure.

For each $n \geq M$ set

$$B_n = \{\omega : \pi(A_n^\omega) \leq n - c \ln n\}.$$

Fix $n \geq M$ and let $d = [n - c \ln n]$. For a_0, a_1, \dots, a_d and $r \leq d$ set $A_r^a = [a_0, a_1, \dots, a_{r-1}]$ and $D_{r,n}^a = \{\omega = (\omega_0, \omega_1, \dots) : \omega_k = a_{k-r[k/r]} \forall k = r, r+1, \dots, n-1\}$. Then by (2.1), (2.2) and Lemma 3.1,

$$\begin{aligned} \mathbb{P}(B_n) & \leq \sum_{r=1}^d \mathbb{P}\{\omega : A_n^\omega \cap T^{-r}(A_n^\omega) \neq \emptyset\} = \sum_{r=1}^d \left(\sum_{a_0, \dots, a_{r-1} \in \mathcal{A}} \mathbb{P}(A_r^a \cap D_{r,n}^a) \right) \\ & \leq (1 + \psi_0) \sum_{r=1}^d \sum_{a_0, \dots, a_{r-1} \in \mathcal{A}} \mathbb{P}(A_r^a) \mathbb{P}(D_{r,n}^a) \leq (1 + \psi_0) \sum_{r=1}^d e^{-\Gamma(n-r)} \\ & \times \sum_{a_0, \dots, a_{r-1} \in \mathcal{A}} \mathbb{P}(A_r^a) = (1 + \psi_0) \sum_{r=1}^d e^{-\Gamma(n-r)} \leq d(1 + \psi_0) e^{-\Gamma(n-d)} \\ & \leq n(1 + \psi_0) e^{-\Gamma c \ln n} = (1 + \psi_0) n^{1-\Gamma c} = (1 + \psi_0) n^{-2}. \end{aligned}$$

It follows that

$$\sum_{n=M}^{\infty} \mathbb{P}(B_n) \leq (1 + \psi_0) \sum_{n=M}^{\infty} n^{-2} < \infty.$$

Now from the Borel-Cantelli lemma we obtain that $\mathbb{P}\{B_n \text{ i.o.}\} = 0$ where i.o. stands for "infinitely often". Set $D = \Omega \setminus \Omega_{\mathbb{P}}$ which, recall, is the union of cylinders A with $\mathbb{P}(A) = 0$. Since $\mathbb{P}(D) = 0$ then $\mathbb{P}(\Omega \setminus \Omega^*) = \mathbb{P}(D \cup \{B_n \text{ i.o.}\}) = 0$ completing the proof of the corollary. \square

5. COMPOUND POISSON APPROXIMATION

5.1. Proof of Theorem 2.3. First, recall that assertions concerning $\rho = \rho_A$ are contained in Lemmas 3.7 and 3.8. Throughout this subsection $A \in \mathcal{C}_n$ will be fixed, and so we will write X_k and S_N for X_k^A and S_N^A , respectively. Next, set $K = 5d_\ell rn$ and $\hat{X}_\alpha = X_\alpha$ if $K < \alpha \leq N$ and $\hat{X}_\alpha = 0$ if $\alpha \leq K$ or $\alpha > N$. Now define

$$U = \sum_{\alpha=1}^N \hat{X}_\alpha = \sum_{\alpha=K+1}^N X_\alpha \text{ and } X_{\alpha,j} = (1 - \hat{X}_{\alpha-\kappa})(1 - \hat{X}_{\alpha+j\kappa}) \prod_{k=0}^{j-1} \hat{X}_{\alpha+k\kappa}.$$

Observe that for any m ,

$$(5.1) \quad |S_N - U| \leq \sum_{\alpha=1}^K X_\alpha \text{ and } |U - \sum_{\alpha=K+1}^N \sum_{j=1}^m j X_{\alpha,j}| \leq m \sum_{\alpha=K+1}^N \prod_{k=0}^{m-1} X_{\alpha+k\kappa}$$

with κ defined by (2.11). Introduce also

$$I_0 = \{K+1, \dots, N\} \times \{1, \dots, n_0\},$$

where, recall, $n_0 = [\frac{n}{r}]$,

$$\lambda_{\alpha,j} = EX_{\alpha,j}, \lambda = \sum_{(\alpha,j) \in I_0} \lambda_{\alpha,j} \text{ and } s = t(1 - \rho)$$

with ρ defined in (2.11).

Next, we estimate $|\lambda - s|$. For each $i = 1, 2, \dots, \ell$ set for brevity $c_i = n + d_i\kappa$. Then

$$\begin{aligned} |\lambda - s| &\leq (\mathbb{P}(A))^\ell + \left| \sum_{\alpha=K+1}^N \sum_{j=1}^{n_0} \lambda_{\alpha,j} - N(\mathbb{P}(A))^\ell(1 - \rho) \right| \\ &\leq (K+1)(\mathbb{P}(A))^\ell + \sum_{\alpha=K+1}^N \left| \sum_{j=1}^{n_0} \lambda_{\alpha,j} - (\mathbb{P}(A))^\ell(1 - \rho) \right| \\ &= (K+1)(\mathbb{P}(A))^\ell + \sum_{\alpha=K+1}^N \left| \mathbb{P}\left(\bigcup_{j=1}^{n_0} \{X_{\alpha,j} = 1\}\right) - (\mathbb{P}(A))^\ell + \prod_{i=1}^{\ell} \mathbb{P}(R^{c_i/r}) \right| \\ &\leq 2Ke^{-\Gamma n} + \sum_{\alpha=K+1}^N \mathbb{P}\left\{ \prod_{k=0}^{n_0} X_{\alpha+k\kappa} = 1 \right\} \\ &\quad + \sum_{\alpha=K+1}^N \left| \mathbb{P}\{(1 - X_{\alpha-\kappa})X_\alpha = 1\} - \mathbb{P}\{X_\alpha = 1\} + \mathbb{P}\left(\bigcap_{i=1}^{\ell} T^{-d_i(\alpha-\kappa)} R^{c_i/r}\right) \right| \\ &\quad + \sum_{\alpha=K+1}^N \left(|\mathbb{P}\{X_\alpha = 1\} - (\mathbb{P}(A))^\ell| + \left| \mathbb{P}\left(\bigcap_{i=1}^{\ell} T^{-d_i(\alpha-\kappa)} R^{c_i/r}\right) - \prod_{i=1}^{\ell} \mathbb{P}(R^{c_i/r}) \right| \right) \\ &= 2Ke^{-\Gamma n} + \sigma_1 + \sigma_2 + \sigma_3. \end{aligned}$$

Here σ_1, σ_2 and σ_3 denote the first, second and third sums, respectively, and we use in the last inequality above that

$$\sum_{\alpha=K+1}^N \mathbb{P}\left\{ \prod_{k=0}^{n_0} X_{\alpha+k\kappa} = 1 \right\} \geq \sum_{\alpha=K+1}^N \left| \mathbb{P}\{(1 - X_{\alpha-\kappa})X_\alpha = 1\} - \mathbb{P}\left(\bigcup_{j=1}^{n_0} \{X_\alpha = 1\}\right) \right|.$$

In order to estimate σ_1 we observe that the choice of K gives $\alpha(d_{i+1} - d_i) \geq 3d_i n$ for any $\alpha > K$ and $i = 1, \dots, \ell - 1$. It follows that whenever $0 \leq m \leq n_0$ and $\alpha > K$ there exist disjoint sets of integers Q_1, Q_2, \dots, Q_ℓ satisfying (3.2) with $k \geq 1$ and such that $T^{-d_i\alpha} \cap_{l=0}^{n_0} T^{-d_il\kappa} A \in \mathcal{F}_{Q_i}$, $i = 1, \dots, \ell$. Since r divides $d_i\kappa$ and $d_i\kappa \leq d_i r < n$ by the assumption then for such m each $\cap_{l=0}^m T^{-d_il\kappa} A$ is contained in $D_{i,m} \cap T^{-d_im\kappa} A$ where $D_{i,m}$ is a cylinder set of the length $d_im\kappa \geq rm$ and such that $D_{i,m} \in \mathcal{F}_{0,d_im\kappa-1}$ while, clearly, $T^{-d_im\kappa} A \in \mathcal{F}_{d_im\kappa, \infty}$. Hence, relying on Lemmas 3.1 and 3.2 we conclude from here that

$$\begin{aligned} (5.2) \quad \sum_{\alpha=K+1}^N \mathbb{P}\left\{ \prod_{l=0}^m X_{\alpha+l\kappa} = 1 \right\} &= \sum_{\alpha=K+1}^N \mathbb{P}\left(\bigcap_{i=1}^{\ell} T^{-d_i\alpha} \cap_{l=0}^m T^{-d_il\kappa} A \right) \\ &\leq (1 + \psi_0)^\ell \sum_{\alpha=K+1}^N \prod_{i=1}^{\ell} \mathbb{P}\left(\cap_{l=0}^m T^{-d_il\kappa} A \right) \\ &\leq (1 + \psi_0)^{2\ell} N(\mathbb{P}(A))^\ell e^{-\Gamma \ell rm} \leq (1 + \psi_0)^{2\ell} te^{-\Gamma \ell rm}. \end{aligned}$$

In particular, taking $m = n_0$ we obtain

$$(5.3) \quad \sigma_1 \leq (1 + \psi_0)^{2\ell} te^{-\Gamma \ell rn_0} \leq (1 + \psi_0)^{2\ell} te^{-\Gamma n/2}.$$

Next we show that the term σ_2 vanishes. Since r divides $d_i\kappa$ for each $i = 1, \dots, \ell$ we have that $R^{c_i/r} = R^{d_i\kappa/r} \cap (T^{-d_i\kappa}A)$. It follows that for any $\alpha > K$,

$$\begin{aligned} \bigcap_{i=1}^{\ell} T^{-d_i(\alpha-\kappa)} R^{c_i/r} &= \left(\bigcap_{i=1}^{\ell} T^{-d_i(\alpha-\kappa)} R^{d_i\kappa/r} \right) \cap \left(\bigcap_{i=1}^{\ell} T^{-d_i\alpha} A \right) \\ &= \left(\bigcap_{i=1}^{\ell} T^{-d_i(\alpha-\kappa)} A \right) \cap \left(\bigcap_{i=1}^{\ell} T^{-d_i\alpha} A \right) = \{X_{\alpha-\kappa} X_{\alpha} = 1\}. \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{P}\{X_{\alpha} = 1\} - \mathbb{P}\left(\bigcap_{i=1}^m T^{-d_i(\alpha-\kappa)} R^{c_i/r}\right) &= \mathbb{P}\{X_{\alpha} = 1\} - \mathbb{P}\{X_{\alpha-\kappa} X_{\alpha} = 1\} \\ &= \mathbb{P}\{(1 - X_{\alpha-\kappa}) X_{\alpha} = 1\}, \end{aligned}$$

and so $\sigma_2 = 0$.

Next, we estimate σ_3 using Lemma 3.2 similarly to above which gives

$$(5.4) \quad \sigma_3 \leq 2^{\ell} N \psi_n (\mathbb{P}(A))^{\ell} \leq 2^{\ell} t \psi_n.$$

Hence, by (5.3) and (5.4),

$$\begin{aligned} (5.5) \quad |\lambda - s| &\leq 2K e^{-\Gamma n} + (1 + \psi_0)^{2\ell} t e^{-\frac{\Gamma}{2}n} + 2^{\ell} t \psi_n \\ &\leq 2(1 + \psi_0)^{2\ell} K(t+1) e^{-\frac{\Gamma}{2}n} + 2^{\ell} t \psi_n. \end{aligned}$$

Next, assume that $\lambda = 0$. Then by (5.1),

$$(5.6) \quad \mathbb{P}\{S_N \neq 0\} \leq \sum_{\alpha=1}^K \mathbb{P}\{X_{\alpha} = 1\} + \sigma_1 \leq K \mathbb{P}(A) + \sigma_1.$$

Let η_1, η_2, \dots be a sequence of i.i.d. random variables with $\mathbb{P}\{\eta_1 \in \mathbb{N}^+\} = 1$ independent of a Poisson random variable W with the parameter s and $Z = \sum_{k=1}^W \eta_k$. Then by (5.3), (5.5) and (5.6) for any $L \subset \mathbb{N}$,

$$\begin{aligned} (5.7) \quad |\mathbb{P}\{S_N \in L\} - \mathbb{P}\{Z \in L\}| &\leq \mathbb{P}\{S_N \neq 0\} + |\mathbb{P}\{S_N = 0\} - \mathbb{P}\{Z = 0\}| \\ &+ \mathbb{P}\{Z \neq 0\} \leq 2(\mathbb{P}\{S_N \neq 0\} + \mathbb{P}\{Z \neq 0\}) = 2(\mathbb{P}\{S_N \neq 0\} + (1 - e^{-s})) \\ &\leq 2(\mathbb{P}\{S_N \neq 0\} + s) \leq 8(1 + \psi_0)^{2\ell} K(t+1) e^{-\frac{\Gamma}{2}n} + 2^{\ell+1} t \psi_n. \end{aligned}$$

Hence, if $\lambda = 0$ the theorem follows for any such i.i.d. sequence η_1, η_2, \dots , and so we can assume that $\lambda > 0$.

Define

$$I = \{(\alpha, j) \in I_0 : \mathbb{P}\{X_{\alpha,j} = 1\} > 0\}$$

When $\lambda > 0$ then $I \neq \emptyset$. For each $j \in \{1, \dots, n_0\}$ set

$$\lambda_j = \lambda^{-1} \sum_{\alpha=K+1}^N \lambda_{\alpha,j}$$

We choose an i.i.d. sequence $\{\eta_k\}_{k=1}^{\infty}$ such that $\mathbb{P}\{\eta_1 = j\} = \lambda_j$ for each $1 \leq j \leq n_0$ and set, again, $Z = \sum_{k=1}^W \eta_k$ where, as before, W is a Poisson random with the parameter s independent of η_k 's. Set $\mathbf{X} = \{X_{\alpha,j}\}_{(\alpha,j) \in I}$ and let $\mathbf{Y} = \{Y_{\alpha,j}\}_{(\alpha,j) \in I}$ be a collection of independent random variables such that each $Y_{\alpha,j}$ has the Poisson distribution with the parameter $\lambda_{\alpha,j}$. Given $(a_{\alpha,j})_{(\alpha,j) \in I} = a \in \mathbb{N}^I$ define

$$f(a) = \sum_{(\alpha,j) \in I} ja_{\alpha,j} \text{ and } h_L(a) = \mathbb{I}_{f(a) \in L}.$$

Then

$$(5.8) \quad \begin{aligned} |\mathbb{P}\{S_N \in L\} - \mathbb{P}\{Z \in L\}| &\leq |\mathbb{P}\{S_N \in L\} - Eh_L(\mathbf{X})| \\ &+ |Eh_L(\mathbf{X}) - Eh_L(\mathbf{Y})| + |Eh_L(\mathbf{Y}) - \mathbb{P}\{Z \in L\}| = \delta_1 + \delta_2 + \delta_3 \end{aligned}$$

where δ_1, δ_2 and δ_3 denote the respective terms in the right hand side of (5.8).

By (5.1) and (5.3) we obtain that

$$(5.9) \quad \begin{aligned} \delta_1 &\leq 2\mathbb{P}\{S_N \neq f(\mathbf{X})\} \leq 2 \sum_{\alpha=1}^K \mathbb{P}\{X_\alpha = 1\} + 2\sigma_1 \leq 2K\mathbb{P}(A) + 2\sigma_1 \\ &\leq 2K\mathbb{P}(A) + 2(1 + \psi_0)^{2\ell} te^{-\frac{\Gamma}{2}n} \leq 4(1 + \psi_0)^{2\ell} K(t + 1)e^{-\frac{\Gamma}{2}n}. \end{aligned}$$

In order to estimate δ_2 we use Theorem 2 from [6]. Note that by the definition of I for each $(\alpha, j) \in I$ we have $\mathbb{P}\{X_{\alpha,j} = 1\} > 0$, and so the use of the theorem is justified. For each $(\alpha, j) \in I$ define

$$B_{\alpha,j} = \{(\beta, k) \in I : \exists i_1, i_2 = 1, \dots, \ell \text{ such that } |d_{i_1}\alpha - d_{i_2}\beta| < K\}.$$

By Theorem 2 in [6] we see that

$$(5.10) \quad \delta_2 \leq \|\mathcal{L}(f(\mathbf{X})) - \mathcal{L}(f(\mathbf{Y}))\| \leq 2(2b_1 + 2b_2 + b_3),$$

where

$$\|\mathcal{L}(\xi) - \mathcal{L}(\zeta)\| = 2 \sup_{L \subset \mathbb{N}} |\mathbb{P}\{\xi \in L\} - \mathbb{P}\{\zeta \in L\}|$$

is the total variation distance between distributions of nonnegative integer valued random variables ξ and ζ ,

$$\begin{aligned} b_1 &= \sum_{(\alpha, j) \in I} \left(\sum_{(\beta, k) \in B_{\alpha, j}} \mathbb{P}\{X_{\alpha, j} = 1\} \mathbb{P}\{X_{\beta, k} = 1\} \right), \\ b_2 &= \sum_{(\alpha, j) \in I} \left(\sum_{(\alpha, j) \neq (\beta, k) \in B_{\alpha, j}} \mathbb{P}\{X_{\alpha, j} = 1, X_{\beta, k} = 1\} \right) \text{ and} \\ b_3 &= \sum_{(\alpha, j) \in I} E|E(X_{\alpha, j} - \lambda_{\alpha, j} | \mathcal{B}_{\alpha, j})| \end{aligned}$$

where $\mathcal{B}_{\alpha, j} = \sigma\{X_{\beta, k} : (\beta, k) \notin B_{\alpha, j}\}$. For each $(\alpha, j) \in I$ it follows by Lemma 3.2 that

$$\mathbb{P}\{X_{\alpha, j} = 1\} \leq \mathbb{P}\{X_\alpha = 1\} \leq 2(\mathbb{P}(A))^\ell.$$

Since the number of elements in $B_{\alpha, j}$ and in I do not exceed $2K\ell^2 n$ and nN , respectively, we obtain from here that

$$(5.11) \quad b_1 \leq 8n^2 N K \ell^2 (\mathbb{P}(A))^{2\ell} \leq 8K\ell^2 tn^2 (\mathbb{P}(A))^\ell \leq 8K\ell^2 tn^2 e^{-\Gamma n}.$$

Now we estimate b_2 . Fix $(\alpha, j) \in I$, let $(\alpha, j) \neq (\beta, k) \in B_{\alpha, j}$ and set $F = \{X_{\alpha, j} = 1, X_{\beta, k} = 1\}$. We want to estimate $\mathbb{P}(F)$ from above. Clearly, if $\alpha = \beta$ then $F = \emptyset$, so without loss of generality we can assume that $\alpha < \beta$. Suppose, first, that $\alpha + \frac{n-3r}{d_\ell} > \beta$ and show that in this case $F = \emptyset$. Indeed, assume by contradiction that $F \neq \emptyset$ then by Lemma 3.6 it follows that $\alpha \leq \beta - \kappa$. Let $\omega \in F$ then $X_{\beta, k}(\omega) = 1$, and so $X_{\beta-\kappa}(\omega) = 0$ and $X_\beta(\omega) = 1$. Hence, there exists an $1 \leq i_0 \leq \ell$ such that

$$\mathbb{I}_A \circ T^{d_{i_0}(\beta-\kappa)}(\omega) = 0 \text{ and } \mathbb{I}_A \circ T^{d_{i_0}\beta}(\omega) = 1.$$

It follows that for $c = d_{i_0}\kappa$,

$$(5.12) \quad \mathbb{I}_{R^{c/r}} \circ T^{d_{i_0}(\beta-\kappa)}(\omega) = 0.$$

By our assumption,

$$d_{i_0}(\beta - \alpha) < d_\ell(\beta - \alpha) < n - 3r < (n_0 - 2)r.$$

Write $d_{i_0}(\beta - \alpha) = ur + v$ where $u, v \in \mathbb{N}$ and $v < r$. Then $d_{i_0}\alpha + ur \leq d_{i_0}\alpha + (n_0 - 2)r$. Since $X_{\alpha,j}(\omega) = 1$, and so $\mathbb{I}_A \circ T^{d_{i_0}\alpha}(\omega) = 1$, we obtain from the last inequality and the definition of n_0 that $\mathbb{I}_{R^2} \circ T^{d_{i_0}\alpha+ur}(\omega) = 1$ where $R^2 = R^{2r/r}$ is the concatenation of two copies of R . Since $X_{\beta,k}(\omega) = 1$, and so $\mathbb{I}_A \circ T^{d_{i_0}\beta}(\omega) = 1$, we obtain also that

$$\mathbb{I}_{R^2} \circ T^{d_{i_0}\alpha+ur+v}(\omega) = \mathbb{I}_{R^2} \circ T^{d_{i_0}\beta}(\omega) = 1.$$

Hence, $R^2 \cap T^{-v}(R^2) \neq \emptyset$. By the assumption $v < r$ and if $v > 0$ then $\pi(R) \leq \pi(R^2) \leq v < r$. If r is not divisible by $\pi(R)$ then by Lemma 3.5 we would have $\pi(R^2) = r$ contradicting the above inequality and if $\pi(R)$ divides r then $\pi(R) \in \mathcal{O}(A)$ contradicting $\pi(A) = r$. Hence, $v = 0$, and so $d_{i_0}(\beta - \alpha) = ur$. Since $X_{\alpha,j}(\omega) = 1$, r divides $d_{i_0}\kappa$ and $n \geq ur = d_{i_0}(\beta - \alpha) \geq d_{i_0}\kappa$ it follows from here that

$$\begin{aligned} 1 &= X_\alpha(\omega) \leq \mathbb{I}_A \circ T^{d_{i_0}\alpha}(\omega) \leq \mathbb{I}_{R^{ur/r}} \circ T^{d_{i_0}\alpha}(\omega) \\ &= (\mathbb{I}_{R^{(ur-d_{i_0}\kappa)/r}} \circ T^{d_{i_0}\alpha}(\omega)) \cdot (\mathbb{I}_{R^{c/r}} \circ T^{(d_{i_0}\alpha+ur-d_{i_0}\kappa)}(\omega)) \\ &\leq \mathbb{I}_{R^{c/r}} \circ T^{(d_{i_0}\alpha+ur-d_{i_0}\kappa)}(\omega) = \mathbb{I}_{R^{c/r}} \circ T^{d_{i_0}(\beta-\kappa)}(\omega) \end{aligned}$$

where, again, $c = d_{i_0}\kappa$ and we set $R^{0/r} = \Omega$. Hence, $\mathbb{I}_{R^{c/r}} \circ T^{d_{i_0}(\beta-\kappa)}(\omega) = 1$ contradicting (5.12), and so $F = \emptyset$ in this case.

Thus, we can assume that $\alpha + \frac{n-3r}{d_\ell} \leq \beta$, and so $d_\ell\alpha + n \leq d_\ell\beta + 3r$. Hence, by Lemma 3.2,

$$\begin{aligned} \mathbb{P}\{X_{\alpha,j} = 1, X_{\beta,k} = 1\} &\leq \mathbb{P}\{X_\alpha = 1, \mathbb{I}_A \circ T^{d_\ell\beta} = 1\} \\ &\leq \mathbb{P}\{X_\alpha = 1, \mathbb{I}_{R^{(n-3r)/r}} \circ T^{d_\ell\beta+3r} = 1\} \\ &\leq (1 + \psi_0)\mathbb{P}\{X_\alpha = 1\}\mathbb{P}(R^{(n-3r)/r}) \leq 4\psi_0(\mathbb{P}(A))^\ell e^{-\Gamma(n-3r)}. \end{aligned}$$

Since the number of elements in $B_{\alpha,j}$ and in I do not exceed $2K\ell^2n$ and nN , respectively, we obtain that

$$(5.13) \quad b_2 \leq 8n^2NK\ell^2\psi_0(\mathbb{P}(A))^\ell e^{-\Gamma(n-3r)} \leq 8\ell^2\psi_0Ktn^2e^{-\frac{\Gamma}{2}n}.$$

In order to estimate b_3 we use Lemma 3.3 with

$$\begin{aligned} Q = Q_{\alpha,j} &= \{d_i(\alpha + l\kappa) + m : i = 1, \dots, \ell; l = -1, 0, 1, \dots, j, m = 0, 1, \dots, n-1\} \\ \text{and } \tilde{Q} &= \tilde{Q}_{\alpha,j} = \{d_i(\beta + l\kappa) + m : \beta \notin B_{\alpha,j}, \\ &i = 1, \dots, \ell, l = -1, 0, 1, \dots, n_0, m = 0, 1, \dots, n-1\}. \end{aligned}$$

Then by the choice of K the conditions of Lemma 3.3 are satisfied with $d = n$ and such Q and \tilde{Q} . Taking into account that $B_{\alpha,j} \subset \mathcal{F}_{\tilde{Q}}$ we obtain from (2.1), (2.2) and Lemma 3.3 that

$$\begin{aligned} E|E(X_{\alpha,j} - \lambda_{\alpha,j} | \mathcal{B}_{\alpha,j})| &= E|E(E(X_{\alpha,j} - \lambda_{\alpha,j} | \mathcal{F}_{\tilde{Q}}) | \mathcal{B}_{\alpha,j})| \\ &\leq E|E(X_{\alpha,j} - \lambda_{\alpha,j} | \mathcal{F}_{\tilde{Q}})| \leq 2^{2\ell+4}\psi_n\mathbb{P}\{X_{\alpha,j} = 1\}. \end{aligned}$$

For each $1 \leq i \leq \ell$ set $c_i = n + \kappa d_i(j-1)$. Then by Lemma 3.2 we see that

$$\begin{aligned} \mathbb{P}\{X_{\alpha,j} = 1\} &\leq \mathbb{P}\left\{\bigcap_{i=1}^m \mathbb{I}_{R^{c_i/r}} \circ T^{d_i\alpha}\right\} \leq 2 \prod_{i=1}^{\ell} \mathbb{P}(R^{c_i/r}) \\ &\leq 2(1 + \psi_0) \prod_{i=1}^{\ell} (\mathbb{P}(A)e^{-\Gamma\kappa d_i(j-1)}) \leq 2(1 + \psi_0)(\mathbb{P}(A))^\ell e^{-\Gamma\ell r(j-1)}. \end{aligned}$$

It follows from the above estimates that

$$\begin{aligned}
(5.14) \quad b_3 &= \sum_{(\alpha,j) \in I} E|E(X_{\alpha,j} - \lambda_{\alpha,j} | \mathcal{B}_{\alpha,j})| \leq \sum_{\alpha=K+1}^N \sum_{j=1}^{n_0} 2^{2\ell+4} \psi_n \mathbb{P}\{X_{\alpha,j} = 1\} \\
&\leq 2^{2\ell+5} (1 + \psi_0) \psi_n N(\mathbb{P}(A))^\ell \sum_{j=1}^{n_0} e^{-\Gamma(j-1)} \\
&\leq 2^{2\ell+5} (1 + \psi_0) \psi_n N(\mathbb{P}(A))^\ell \frac{1}{1-e^{-\Gamma}} \leq 2^{2\ell+5} (1 + \psi_0) t \psi_n (1 - e^{-\Gamma})^{-1}.
\end{aligned}$$

Next, we estimate δ_3 . Given a random variable ξ we denote by φ_ξ the characteristic function of ξ . Let ξ be a Poisson random variable with a parameter λ independent of $\{\eta_k\}_{k=1}^\infty$ and set $\Psi = \sum_{k=1}^\xi \eta_k$. Then for each $s \in \mathbb{R}$,

$$\varphi_\Psi(s) = \sum_{l=0}^\infty \mathbb{P}\{\xi = l\} \prod_{k=1}^l \varphi_{\eta_k}(s) = e^{-\lambda} \sum_{l=0}^\infty \frac{\lambda^l}{l!} (\varphi_{\eta_1}(s))^l = \exp(\lambda \sum_{j=1}^{n_0} \lambda_j (e^{ij s} - 1))$$

and

$$\varphi_f(\mathbf{Y})(s) = \prod_{(\alpha,j) \in I} \varphi_{Y_{\alpha,j}}(js) = \exp\left(\sum_{(\alpha,j) \in I} \lambda_{\alpha,j} (e^{ij s} - 1)\right) = \exp\left(\lambda \sum_{j=1}^{n_0} \lambda_j (e^{ij s} - 1)\right)$$

warning the reader that the last two formulas are the only places in this paper where i stands for $\sqrt{-1}$ and not for an integer. It follows from here that $f(\mathbf{Y})$ and Ψ have the same distribution, and so by Lemmas 3.2 and 3.4,

$$\begin{aligned}
(5.15) \quad \delta_3 &= |\mathbb{P}\{\Psi \in L\} - \mathbb{P}\{Z \in L\}| = \sum_{l=0}^\infty |\mathbb{P}\{\xi = l\} - \mathbb{P}\{W = l\}| \mathbb{P}\{\sum_{k=1}^l \eta_k \in L\} \\
&\leq \sum_{l=0}^\infty |\mathbb{P}\{\xi = l\} - \mathbb{P}\{W = l\}| \leq 2e^{|\lambda-s|} |\lambda - s| = 2\wp(|\lambda - s|).
\end{aligned}$$

Finally, (2.14) follows from (5.5), (5.7)–(5.11) and (5.13)–(5.15) completing the proof of Theorem 2.3. \square

5.2. Proof of Corollary 2.4. Let $\omega \in \Omega_{\mathbb{P}}$ be a nonperiodic sequence, $A = A_n^\omega$ and $r = r_n^\omega = \pi(A_n^\omega)$. Assume first that $n > r_n^\omega(d_\ell + 6)$ so that Theorem 2.3 can be applied. Then by (5.2) and the definition of the sequence η_1, η_2, \dots in the proof of Theorem 2.3 we obtain that

$$(5.16) \quad \mathbb{P}\{\eta_1 = j\} = \lambda_j = \lambda^{-1} \sum_{\alpha=K+1}^N \lambda_{\alpha,j} \leq \lambda^{-1} (1 + \psi_0)^{\ell+1} t e^{-\Gamma \ell r(j-1)}.$$

Let Ξ be a Poisson random variable with the parameter t and $Z = \sum_{l=1}^W \eta_l$ be the compound Poisson random variable constructed by Theorem 2.3 with $A = A_n^\omega$, $\rho = \rho_n^\omega$ and $r = r_n^\omega$. Then for any $L \subset \mathbb{N}$,

$$\begin{aligned}
(5.17) \quad |\mathbb{P}\{\Xi \in L\} - \mathbb{P}\{Z \in L\}| &\leq \sum_{l=0}^\infty |\mathbb{P}\{\Xi = l\} - \mathbb{P}\{W = l\}| \mathbb{P}\{\sum_{k=1}^l \eta_k \in L\} \\
&\quad + \sum_{l=0}^\infty \mathbb{P}\{\Xi = l\} |\mathbb{P}\{\sum_{k=1}^l \eta_k \in L\} - \mathbb{I}_L(l)|
\end{aligned}$$

where $\mathbb{I}_L(l) = 1$ if $l \in L$ and $= 0$, otherwise.

By Lemma 3.4,

$$(5.18) \quad \sum_{l=0}^\infty |\mathbb{P}\{\Xi = l\} - \mathbb{P}\{W = l\}| \leq 2t\rho_n^\omega e^{t\rho_n^\omega} = 2\wp(t\rho_n^\omega).$$

Next, by (5.16),

$$(5.19) \quad \begin{aligned} |\mathbb{P}\{\sum_{k=1}^l \eta_k \in L\} - \mathbb{I}_L(l)| &\leq 2\mathbb{P}\{\sum_{k=1}^l \eta_k \neq l\} \leq 2l\mathbb{P}\{\eta_1 \neq 1\} \\ &= 2l \sum_{j=2}^{n_0} \lambda_j \leq 2l\lambda^{-1}(1+\psi_0)^{2\ell}te^{-\Gamma\ell r_n^\omega}(1-e^{-\Gamma\ell r_n^\omega})^{-1}. \end{aligned}$$

Now (5.17)–(5.19) yield

$$(5.20) \quad |\mathbb{P}\{\Xi \in L\} - \mathbb{P}\{Z \in L\}| \leq 2\wp(t\rho_n^\omega) + 2\lambda^{-1}(1+\psi_0)^{\ell+1}t^2e^{-\Gamma\ell r_n^\omega}(1-e^{-\Gamma\ell r_n^\omega})^{-1}$$

where we used that $t = E\Xi = e^{-t} \sum_{l=0}^{\infty} l \frac{t^l}{l!}$.

If $n \leq r_n^\omega(d_\ell + 6)$ then we apply Theorem 2.1, and so we can write

$$|\mathbb{P}\{S_N^{A_n^\omega} \in L\} - \mathbb{P}\{\Xi \in L\}| \leq \max(\varepsilon_1(n) + \varepsilon_2(n), \varepsilon_3(n))$$

where $\varepsilon_1(n)$ and $\varepsilon_2(n)$ are right hand sides of (2.14) and (5.20), respectively, while $\varepsilon_3(n) = 16e^{-\Gamma n/(d_\ell+6)}(\ell^2nt + \gamma(n)(1+t^{-1}) + tnl^2(1+\psi_0)) + 2\wp(2^\ell t\psi_n + \gamma(n)e^{-\Gamma n})$.

Clearly, $\varepsilon_1(n), \varepsilon_3(n) \rightarrow 0$ as $n \rightarrow \infty$ and since $\rho_n^\omega \rightarrow 0$ and $r_n^\omega \rightarrow \infty$ as $n \rightarrow \infty$ by Lemma 3.7 then $\varepsilon_2(n) \rightarrow 0$ as $n \rightarrow \infty$, as well, and so the assertion of Corollary 2.4 follows. \square

5.3. Proof of Corollary 2.6. Let $t > 0$ be given. If $n > r(d_\ell + 6)$ then we take W and Z to be as in Theorem 2.3. Note that $\mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > t\} = \mathbb{P}\{S_N^A = 0\}$ and $\mathbb{P}\{Z = 0\} = \mathbb{P}\{W = 0\}$, and so by Theorem 2.3,

$$(5.21) \quad \begin{aligned} |\mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > t\} - \mathbb{P}\{W = 0\}| &= |\mathbb{P}\{S_N^A = 0\} - \mathbb{P}\{Z = 0\}| \\ &\leq 2^{2\ell+7}(1+\psi_0)^{2\ell}(t+1)(d_\ell\ell^2n^4e^{-\Gamma n/2} + \psi_n(1-e^{-\Gamma})^{-1}) \\ &\quad + 2\wp(2^\ell t\psi_n + 10e^{-\Gamma n/2}(1+\psi_0)^{2\ell}d_\ell n^2(t+1)) \end{aligned}$$

On the other hand, if $n \leq r(d_\ell + 6)$ then by Theorem 2.1 (with q_i 's being linear),

$$(5.22) \quad |\mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > t\} - P_s\{0\}| \leq |\mathbb{P}\{S_N^A = 0\} - P_t\{0\}| + |P_t\{0\} - P_s\{0\}|.$$

where $s = t(1-\rho)$. Furthermore, $\kappa \geq \frac{r}{d_1}$, and so

$$(5.23) \quad \begin{aligned} \rho &= \prod_{i=1}^{\ell} \mathbb{P}\{R^{(n+d_i\kappa)/r} \mid A\} \leq \frac{\mathbb{P}(R^{(n+d_1\kappa)/r})}{\mathbb{P}(A)} \leq \psi_0 \mathbb{P}(R^{(d_1\kappa)/r}) \\ &\leq \psi_0 e^{-\Gamma d_1 \kappa} \leq \psi_0 e^{-\Gamma r}. \end{aligned}$$

By Theorem 2.1,

$$(5.24) \quad \begin{aligned} |\mathbb{P}\{S_N^A = 0\} - P_t\{0\}| &\leq 16e^{-\Gamma n}(\ell^2nt + \gamma(n)(1+t^{-1})) \\ &\quad + 6(1+\psi_0)tnl^2e^{-\Gamma r} + 2\wp(2^\ell t\psi_n + \gamma(n)e^{-\Gamma n}) \end{aligned}$$

and by (5.23),

$$(5.25) \quad |P_t\{0\} - P_s\{0\}| \leq |t-s| \leq t\psi_0 e^{-\Gamma r}.$$

Taking into account that $\gamma(n) \leq 2n$ when q_i 's are as in Theorem 2.3 and that $r \geq n/(d_\ell + 6)$ in (5.24) we obtain the estimate of Corollary 2.6 from (5.21)–(5.25). \square

5.4. Proof of Corollary 2.7. First, observe that our mixing conditions imply, in particular, that \mathbb{P} is ergodic. Set $\iota = \sup_{A \in \mathcal{C}_n, n \geq 1} \rho_A$ and $b(n) = \frac{2 \ln n}{1-\iota}$ recalling that $\iota < 1$ by (2.12). Now, by Corollary 2.6 for any $A \in \mathcal{C}_n$ with $\mathbb{P}(A) > 0$,

$$(5.26) \quad \begin{aligned} \mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > b(n)\} &\leq e^{-(1-\iota)b(n)} \\ &+ 2^{2\ell+8}(1+\psi_0)^{2\ell}(b(n)+1)(d_\ell \ell^2 n^4(1+\frac{1}{b(n)})e^{-\Gamma n/(d_\ell+6)} + \psi_n(1-e^{-\Gamma})^{-1}) \\ &+ 2\wp(2^\ell b(n)\psi_n + 10e^{-\Gamma n/2}(1+\psi_0)^{2\ell} d_\ell n^2(b(n)+1)) \end{aligned}$$

and

$$(5.27) \quad \begin{aligned} \mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A \leq n^{-2}\} &= 1 - \mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > n^{-2}\} \\ &\leq |1 - e^{-(1-\rho_A)n^{-2}}| + |\mathbb{P}\{(\mathbb{P}(A))^\ell \tau_A > n^{-2}\} - e^{-(1-\rho_A)n^{-2}}| \\ &\leq n^{-2} + 2^{2\ell+8}(1+\psi_0)^{2\ell}(n^{-2}+1)(d_\ell \ell^2 n^4(1+n^2)e^{-\Gamma n/(d_\ell+6)} \\ &+ \psi_n(1-e^{-\Gamma})^{-1}) + 2\wp(2^\ell n^{-2}\psi_n + 10e^{-\Gamma n/2}(1+\psi_0)^{2\ell} d_\ell(1+n^2)) \end{aligned}$$

. Applying (5.26) and (5.27) to $A = A_n^\omega$ with $\omega \in \Omega_{\mathbb{P}}$ we obtain by the Borel-Cantelli lemma that there exists a random variable $m = m(\varpi)$ on Ω finite \mathbb{P} -a.s. and such that for all $n \geq m(\varpi)$,

$$(5.28) \quad n^{-2} < (\mathbb{P}(A_n^\omega))^\ell \tau_{A_n^\omega}(\varpi) \leq b(n)$$

which implies (2.18). Finally, if (2.19) holds true then (2.20) follows from (2.18) and the Shannon-McMillan-Breiman theorem (see, for instance, [21]). \square

6. I.I.D. CASE

6.1. Proof of Theorem 2.8. We use the same notation as in the proof of Theorem 2.3 and as there we can assume that $\lambda > 0$ and $N > K = 5d_\ell r n$. In order to derive (2.21) we will estimate $|\mathbb{P}\{Z \in L\} - \mathbb{P}\{Y \in L\}|$ for any $L \subset \mathbb{N}$ and combine it with (2.14). Observe that under the assumption that the coordinate projections from Ω onto \mathcal{A} are i.i.d. random variables it follows that

$$(6.1) \quad \rho = \prod_{i=1}^{\ell} \mathbb{P}\{R^{(n+d_i\kappa)/r} \mid A\} = \prod_{i=1}^{\ell} \mathbb{P}(R^{(d_i\kappa)/r}) = (\mathbb{P}(R))^{k_0} \leq \mathbb{P}(R),$$

where $k_0 = \frac{\kappa}{r} \sum_{i=1}^{\ell} d_i$, and also that $\lambda_{\alpha,j} = (1-\rho)^2(\mathbb{P}(A))^\ell \rho^{j-1}$ for each $(\alpha, j) \in I$. Let η_1, η_2, \dots be i.i.d. random variables constructed in the proof of Theorem 2.3. Then, for each $j = 1, 2, \dots, n_0 = [n/r]$,

$$\begin{aligned} \mathbb{P}\{\eta_1 = j\} &= \lambda_j = ((N-K) \sum_{l=1}^{n_0} \lambda_{N,l})^{-1} (N-K) \lambda_{N,j} \\ &= ((1-\rho)^2(\mathbb{P}(A))^\ell \sum_{l=1}^{n_0} \rho^{l-1})^{-1} (1-\rho)^2(\mathbb{P}(A))^\ell \rho^{j-1} \\ &= ((1-\rho) \sum_{l=1}^{n_0} \rho^{l-1})^{-1} \mathbb{P}\{\zeta_1 = j\} = \mathbb{P}\{\zeta_1 = j \mid \zeta_1 \in \{1, \dots, n_0\}\} \end{aligned}$$

where ζ_1, ζ_2, \dots are i.i.d. random variables described in the statement of Theorem 2.8. Furthermore, for any $j > n_0$,

$$\mathbb{P}\{\eta_1 = j\} = 0 = \mathbb{P}\{\zeta_1 = j \mid \zeta_1 \in \{1, \dots, n_0\}\}$$

Hence, by the independence for any $l \geq 1$ and each $j_1, \dots, j_l \in \mathbb{N}^+$,

$$\mathbb{P}\{\eta_1 = j_1, \dots, \eta_l = j_l\} = \mathbb{P}\{\zeta_1 = j_1, \dots, \zeta_l = j_l \mid \zeta_1, \dots, \zeta_l \in \{1, \dots, n_0\}\}.$$

It follows from here that

$$\mathbb{P}\left\{\sum_{k=1}^l \eta_k \in L\right\} = \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L \mid \zeta_1, \dots, \zeta_l \in \{1, \dots, n_0\}\right\}.$$

Introduce the event $\Psi = \{\exists k \leq l, k \geq 1 : \zeta_k \notin \{1, \dots, n_0\}\}$. Then by the law of total probability it follows that

$$\begin{aligned} & |\mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L\right\} - \mathbb{P}\left\{\sum_{k=1}^l \eta_k \in L\right\}| \\ &= |\mathbb{P}\{\zeta_1, \dots, \zeta_l \in \{1, \dots, n_0\}\} \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L \mid \zeta_1, \dots, \zeta_l \in \{1, \dots, n_0\}\right\} \\ &\quad + \mathbb{P}(\Psi) \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L \mid \Psi\right\} - \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L \mid \zeta_1, \dots, \zeta_l \in \{1, \dots, n_0\}\right\}| \\ &= 2\mathbb{P}(\Psi) \leq 2 \sum_{k=1}^l \mathbb{P}\{\zeta_k \notin \{1, \dots, n_0\}\} = 2l\rho^{n_0}. \end{aligned}$$

Set again $s = t(1 - \rho)$. Then

$$\begin{aligned} |\mathbb{P}\{Z \in L\} - \mathbb{P}\{Y \in L\}| &\leq \sum_{l=0}^{\infty} \mathbb{P}\{W = l\} |\mathbb{P}\left\{\sum_{k=1}^l \eta_k \in L\right\} - \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L\right\}| \\ &\leq \sum_{l=1}^{n_0} |\mathbb{P}\left\{\sum_{k=1}^l \eta_k \in L\right\} - \mathbb{P}\left\{\sum_{k=1}^l \zeta_k \in L\right\}| + \mathbb{P}\{W > n_0\} \\ &\leq \sum_{l=1}^{n_0} 2l\rho^{n_0} + e^{-s}(e^s - \sum_{l=0}^{n_0} \frac{s^l}{l!}) \leq 2n_0^2 \rho^{n_0} + \frac{t^{n_0+1}}{(n_0+1)!}. \end{aligned}$$

Taking into account that $\psi_n \equiv 0$ under the independency assumption (2.21) follows from here together with (2.14), (6.1) and Lemma 3.1, completing the proof of Theorem 2.8. \square

6.2. Proof of Theorem 2.9. In the i.i.d. case the assertion (i) can be derived easily studying the asymptotic behavior of the compound Poisson distribution constructed Theorem 2.8 but since (i) follows from the more general result of Corollary 2.4 we will prove now only the assertion (ii). Assume that ω is periodic with a period $d \in \mathbb{N}^+$. From Lemma 3.7 it follows that there exist a positive integer $M \geq 1$ such that $r_n^\omega = d$ for all $n \geq M$. Next, by (6.1) for any $n \geq M$,

$$\rho_n^\omega = (\mathbb{P}([\omega_0, \dots, \omega_{r_n^\omega-1}]))^{k_0} = (\mathbb{P}([\omega_0, \dots, \omega_{d-1}]))^{k_0} = \rho^\omega$$

which does not depend on n . Hence, for all $n \geq \max(M, d(d_\ell + 6))$ we can apply Theorem 2.8 in order to obtain that

$$\begin{aligned} & \sup_{L \subset \mathbb{N}} |\mathbb{P}\{U_n^\omega \in L\} - \mathbb{P}\{Y \in L\}| \\ & \leq 2^{2\ell+8}(t+1)d_\ell \ell^2 n^4 e^{-\Gamma n/2} + 2\wp(10d_\ell n^2(t+1)e^{-\Gamma n/2}) + \frac{t^{[n/d]+1}}{([n/d]+1)!} \end{aligned}$$

where $\Gamma = \min\{-\ln(\mathbb{P}\{\omega_0 = a\}) : a \in \mathcal{A}\}$. The expression on the right hand side of the last inequality tends to 0 as $n \rightarrow \infty$, and so (ii) is proved. \square

7. A NONCONVERGENCE EXAMPLE

We assume here that $\mathcal{A} = \{0, 1\}$ and consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\Omega = \{0, 1\}^{\mathbb{N}}$, \mathcal{F} is the σ -algebra generated by cylinder sets and \mathbb{P} is a probability measure on Ω such that the coordinate projections $\{\omega_j\}_{j=0}^{\infty}$ from Ω onto $\{0, 1\}$ are i.i.d. random variables with $\mathbb{P}\{\omega_0 = 0\} = p_0 = 1 - p_1 = 1 - \mathbb{P}\{\omega_0 = 1\}$ where $p_0, p_1 > 0$, $p_0 + p_1 = 1$ and $p_0 \neq p_1$. As before T will denote the left shift on Ω and we introduce here another map $S : \Omega \rightarrow \Omega$ acting by $(S\omega)_n = \omega_n + \omega_{n+1} \pmod{2}$ for any $n \geq 0$ and $\omega = (\omega_0, \omega_1, \dots)$.

Set $\mathbb{P}_0 = S\mathbb{P}$ which is a probability measure on Ω defined by $\mathbb{P}_0(U) = \mathbb{P}(S^{-1}U)$ for any $U \in \mathcal{F}$. We claim that \mathbb{P}_0 is T -invariant. Indeed, let $A = [a_0, a_1, \dots, a_{n-1}]$ be a cylinder set then

$$(7.1) \quad S^{-1}A = [0, \alpha_1, \dots, \alpha_n] \cup [1, \beta_1, \dots, \beta_n]$$

where $\alpha_i + \beta_i = 1$ for all $i = 1, \dots, n$ and $a_i = \alpha_i + \alpha_{i+1} \pmod{2} = \beta_i + \beta_{i+1} \pmod{2}$ for $i = 0, 1, \dots, n-1$ where $\alpha_0 = 0$ and $\beta_0 = 1$. Similarly,

$$(7.2) \quad S^{-1}T^{-1}A = [0, 0, \alpha_1, \dots, \alpha_n] \cup [1, 0, \alpha_1, \dots, \alpha_n] \cup [1, 1, \beta_1, \dots, \beta_n] \cup [0, 1, \beta_1, \dots, \beta_n].$$

It follows from (7.1) and (7.2) that

$$(7.3) \quad \mathbb{P}_0(T^{-1}A) = \mathbb{P}(S^{-1}T^{-1}A) = \mathbb{P}(S^{-1}A) = \mathbb{P}_0(A).$$

Since (7.3) holds true for all cylinder sets, it remains true for their disjoint unions and since it is preserved under monotone limits we obtain that (7.3) is satisfied for any $A \in \mathcal{F}$ which proves our claim.

Next, we will show that \mathbb{P}_0 has a ψ -mixing coefficient ψ_l equal zero for any $l \geq 1$ while $\psi_0 < \infty$. First, observe that ψ_l can be defined using only cylinder sets saying that ψ_l is the infimum of constants $M \geq 0$ such that for each $n \geq 0$ and any cylinder sets $A \in \mathcal{F}_{0,n}$ and $B \in \mathcal{F}_{n+l+1,\infty}$ (as defined at the beginning of Section 2),

$$(7.4) \quad |\mathbb{P}_0(A \cap B) - \mathbb{P}_0(A)\mathbb{P}_0(B)| \leq M\mathbb{P}_0(A)\mathbb{P}_0(B).$$

Indeed, if (7.4) holds true for such cylinder sets then it remains true for their corresponding disjoint unions and it is preserved under monotone limits which yields that (7.4) is satisfied for any sets $A \in \mathcal{F}_{0,n}$ and $B \in \mathcal{F}_{n+l+1,\infty}$, proving the assertion. Now, if $A \in \mathcal{F}_{0,n}$ and $B \in \mathcal{F}_{n+l+1,\infty}$ are cylinder sets then analysing their preimages under S^{-1} similarly to (7.1) we conclude that $S^{-1}A \in \mathcal{F}_{0,n+1}$ and $S^{-1}B \in \mathcal{F}_{n+l+1,\infty}$. Hence, if $l \geq 1$ then $S^{-1}A$ and $S^{-1}B$ are independent events with respect to the probability \mathbb{P} , and so

$$\mathbb{P}_0(A \cap B) - \mathbb{P}_0(A)\mathbb{P}_0(B) = \mathbb{P}(S^{-1}A \cap S^{-1}B) - \mathbb{P}(S^{-1}A)\mathbb{P}(S^{-1}B) = 0$$

implying that $\psi_l = 0$ in this case.

In order to estimate ψ_0 we observe that if $B = [b_n, b_{n+1}, \dots, b_{n+m-1}] = \{\omega = (\omega_0, \omega_1, \dots) : \omega_i = b_i \text{ when } n \leq i \leq n+m-1\}$ then

$$S^{-1}B = [\gamma_n, \gamma_{n+1}, \dots, \gamma_{n+m}] \cup [\delta_n, \delta_{n+1}, \dots, \delta_{n+m}]$$

where $\gamma_i + \delta_i = 1$ for $i = n, \dots, n+m$ and $\gamma_i + \gamma_{i+1} = \delta_i + \delta_{i+1} = b_i$ for $i = n, \dots, n+m-1$. This together with (7.1) yields that

$$\mathbb{P}_0(A \cap B) = \mathbb{P}(S^{-1}A \cap S^{-1}B) = \prod_{i=0}^n p_{\alpha_i} \prod_{j=n+1}^{n+m} p_{\gamma_j} + \prod_{i=0}^n p_{\beta_i} \prod_{j=n+1}^{n+m} p_{\delta_j},$$

where $\alpha_0 = 0$, $\beta_0 = 1$ and without loss of generality we assume that $\alpha_n = \gamma_n$ and $\beta_n = \delta_n$. On the other hand,

$$\begin{aligned}\mathbb{P}_0(A) &= \mathbb{P}(S^{-1}A) = \prod_{i=0}^n p_{\alpha_i} + \prod_{i=0}^n p_{\beta_i} \text{ and} \\ \mathbb{P}_0(B) &= \mathbb{P}(S^{-1}B) = \prod_{j=n}^{n+m} p_{\gamma_j} + \prod_{j=n}^{n+m} p_{\delta_j}.\end{aligned}$$

It follows that

$$\frac{\mathbb{P}_0(A \cap B)}{\mathbb{P}_0(A)\mathbb{P}_0(B)} \leq 2(p_{\gamma_n}^{-1} + p_{\delta_n}^{-1}) = 2(p_0^{-1} + p_1^{-1}).$$

Hence, $\psi_0 \leq 1 + 2(p_0^{-1} + p_1^{-1}) < \infty$ as required.

Let $1^\infty = \omega \in \Omega$ and $t > 0$. For each $n \geq 1$ let A_n^ω , N_n^ω and $U_n^\omega = S_{N_n^\omega}^{A_n^\omega}$ be as defined in Section 2. We now show that the sequence $\{U_n^\omega\}_{n=1}^\infty$ does not converge in distribution when we take \mathbb{P}_0 as the measure on Ω . For two strings S_1 and S_2 of 0 and 1 digits we denote by $S_1 \cdot S_2$ the concatenation of S_1 with S_2 , and for any integer $k \geq 1$ we denote by S_1^k the concatenation of S_1 with itself k times. For every even $n \geq 1$,

$$\mathbb{P}_0(1^n) = \mathbb{P}([1, 0]^{n/2} \cdot 1]) + \mathbb{P}([0, 1]^{n/2} \cdot 0]) = p_1^{\frac{n}{2}+1} p_0^{\frac{n}{2}} + p_0^{\frac{n}{2}+1} p_1^{\frac{n}{2}} = (p_1 p_0)^{\frac{n}{2}}$$

And for every odd $n \geq 1$,

$$\mathbb{P}_0(1^n) = \mathbb{P}([1, 0]^{(n+1)/2}) + \mathbb{P}([0, 1]^{(n+1)/2}) = 2(p_1 p_0)^{\frac{n+1}{2}}.$$

From this it follows that for every even $n \geq 1$,

$$\rho_{A_n^\omega} = \mathbb{P}_0(1^{n+1} \mid 1^n) = \frac{2(p_1 p_0)^{\frac{n+1}{2}}}{(p_1 p_0)^{\frac{n}{2}}} = 2p_1 p_0$$

while for every odd $n \geq 1$,

$$\rho_{A_n^\omega} = \mathbb{P}_0(1^{n+1} \mid 1^n) = \frac{(p_1 p_0)^{\frac{n+1}{2}}}{2(p_1 p_0)^{\frac{n+1}{2}}} = \frac{1}{2}.$$

Now for each $n \geq 1$ set $\theta_n = \mathbb{P}_0\{U_n^\omega = 0\}$. Then from Theorem 2.3 it follows that

$$\lim_{n \rightarrow \infty} \theta_{2n} = e^{-t(1-2p_1 p_0)} \text{ and } \lim_{n \rightarrow \infty} \theta_{2n+1} = e^{-\frac{t}{2}}.$$

Since $p_0 \neq p_1$ then $1 - 2p_1 p_0 \neq \frac{1}{2}$, and so the sequence $\{U_n^\omega\}_{n=1}^\infty$ does not converge in distribution when we take \mathbb{P}_0 as the measure on Ω .

REFERENCES

- [1] M. Abadi, *Exponential approximation for hitting times in mixing processes*, Math. Phys. Electron J. 7, no.2 (2001).
- [2] M. Abadi, *Sharp error terms and necessary conditions for exponential hitting times in mixing processes*, Ann. Probab. 32 (2004), 243–264.
- [3] M. Abadi and B. Saussol, *Hitting and returning into rare events for all alpha-mixing processes*, Stoch. Proc. Appl., to appear.
- [4] M. Abadi and N. Vergne, *Poisson approximation for search of rare words in DNA sequences*, ALEA–Lat. Amer. J. Probab. Meth. Stat. 4 (2008), 223–244.
- [5] M. Abadi and N. Vergne, *Sharp errors for point-wise Poisson approximations in mixing processes*, Nonlinearity 21 (2008), 2871–2885.

- [6] R. Arratia, L. Goldstein and L. Gordon, *Two moments suffice for Poisson approximations: the Chen–Stein method*, Ann. Probab. 17 (1989), 9–25.
- [7] P. Billingsley, *Probability and Measure*, 3d ed., J. Wiley, New York, 1995.
- [8] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Math. 470, Springer–Verlag, Berlin, 1975.
- [9] R.C. Bradley, *Introduction to Strong Mixing Conditions*, Kendrick Press, Heber City, 2007.
- [10] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. D’Analyse Math. 31 (1977), 204–256.
- [11] H. Furstenberg, *Nonconventional ergodic averages*, Proc. Symp. Pure Math. 50, 43–56 (1990).
- [12] A. Galves and B. Schmitt, *Inequalities for hitting times in mixing dynamical systems*, Random Comput. Dyn. 5 (1997), 337–348.
- [13] L. Heinrich, *Mixing properties and central limit theorem for a class of non-identical piecewise monotonic C^2 -transformations*, Mathematische Nachricht. 181, 185–214 (1996).
- [14] N. Haydn and S. Vaienti, *The limiting distribution and error terms for return times of dynamical systems*, Discr. Cont. Dyn. Sys. A, 10, 589–616 (2004).
- [15] N. Haydn and S. Vaienti, *The compound Poisson distribution and return times in dynamical systems*, Probab. Th. Rel. Fields 144, 517–542 (2009).
- [16] Yu. Kifer, *A nonconventional strong law of large numbers and fractal dimensions of some multiple recurrence sets*, Stoch. Dynam., 12 (2012), 1150023.
- [17] Yu. Kifer, *Nonconventional Poisson limit theorems*, Israel J. Math. (2013), to appear.
- [18] Yu. Kifer and S.R.S. Varadhan, *Nonconventional limit theorems in discrete and continuous time via martingales*, Ann. Probab., to appear.
- [19] Y. Kifer, Y. Peres and B. Weiss, *A dimension gap for continued fractions with independent digits*, Israel J. Math. 134 (2001), 61–76.
- [20] Y. Lacroix, *Possible limit laws for entrance times of an ergodic aperiodic dynamical system*, Israel J. Math. 132 (2002), 253–263.
- [21] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, Cambridge, 1983.
- [22] S. Schbath, *Compound Poisson approximation of word counts in DNA sequences*, ESAIM: Probab. Statist. 1 (1995), 1–16.

INSTITUTE OF MATHEMATICS, THE HEBREW UNIVERSITY, JERUSALEM 91904, ISRAEL
E-mail address: kifer@math.huji.ac.il, ariel.rapaport@mail.huji.ac.il